

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
24. November 2005 (24.11.2005)

PCT

(10) Internationale Veröffentlichungsnummer  
WO 2005/111998 A1

(51) Internationale Patentklassifikation<sup>7</sup>: G10L 11/00  
(21) Internationales Aktenzeichen: PCT/EP2005/005004  
(22) Internationales Anmeldedatum:  
9. Mai 2005 (09.05.2005)  
(25) Einreichungssprache: Deutsch  
(26) Veröffentlichungssprache: Deutsch  
(30) Angaben zur Priorität:  
10 2004 023 436.1 10. Mai 2004 (10.05.2004) DE  
(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): FRAUNHOFER-GESELLSCHAFT  
ZUR FÖRDERUNG DER ANGEWANDTEN

FORSCHUNG E.V. [DE/DE]; Hansastrasse 27c, 80686 München (DE).

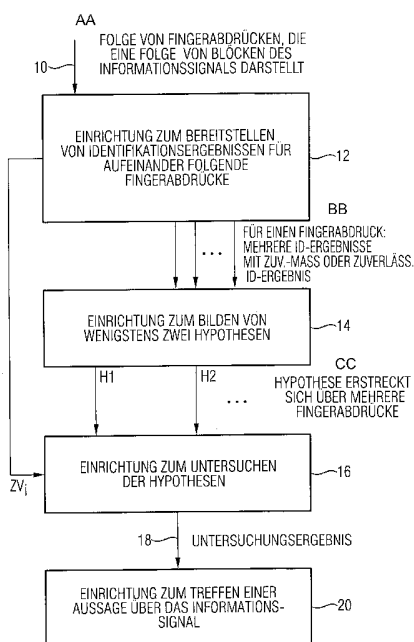
(72) Erfinder; und  
(75) Erfinder/Anmelder (nur für US): HERRE, Jürgen [DE/DE]; Hallerstrasse 24, 91054 Buckenhof (DE). AL-LAMANICHE, Eric [DE/US]; 6400 Christie Avenue, Apt. 3101, Emeryville, CA 94608 (US). HELLMUTH, Oliver [DE/DE]; Geschwister-Vömel-Weg 60, 91052 Erlangen (DE). KASTNER, Thorsten [DE/DE]; Dorfstrasse 17, 96342 Stockheim (DE).

(74) Anwälte: ZINKLER, Franz usw.; Schoppe, Zimmermann, Stöckeler & Zinkler, Postfach 246, 82043 Pullach bei München (DE).

[Fortsetzung auf der nächsten Seite]

(54) Title: DEVICE AND METHOD FOR ANALYZING AN INFORMATION SIGNAL

(54) Bezeichnung: VORRICHTUNG UND VERFAHREN ZUM ANALYSIEREN EINES INFORMATIONSSIGNALS



AA ... SERIES OF FINGERPRINTS THAT DEPICTS A SERIES OF BLOCKS OF THE INFORMATION SIGNAL  
12 ... DEVICE FOR PREPARING IDENTIFICATION RESULTS FOR SUCCESSIVE FINGERPRINTS  
BB ... FOR ONE FINGERPRINT: A NUMBER OF ID RESULTS WITH PERMISSIBLE MASS OR PERMISSIBLE ID RESULT  
14 ... DEVICE FOR FORMING AT LEAST TWO HYPOTHESES  
CC ... HYPOTHESIS COVERING A NUMBER OF FINGERPRINTS  
16 ... DEVICE FOR TESTING THE HYPOTHESES  
18 ... TEST RESULT  
20 ... DEVICE FOR MAKING AN ASSUMPTION CONCERNING THE INFORMATION SIGNAL

(57) Abstract: In order to analyze an information signal, which has a series of blocks of information units, whereby a number of successive blocks of the series of blocks depicts an information entity, identification results for successive fingerprints are prepared (12) while using a series of fingerprints for the series of blocks, whereby an identification result depicts an association of a block of information units with a predetermined information entity. After this, at least two hypotheses are formed (14) from the identification results for the successive fingerprints. A first hypothesis is an assumption for the association of the series of blocks with a first information entity, and the second hypothesis is an assumption for the association of the series of blocks with the second information entity. Afterwards, different hypotheses are tested (16) in order to obtain a test result on the basis of which an assertion concerning the information signal is made (20). This results in obtaining a meaningful and reliable continuous-time analysis of an information signal.

(57) Zusammenfassung: Zum Analysieren eines Informationssignals, das eine Folge von Blöcken von Informationseinheiten aufweist, wobei eine Mehrzahl von aufeinanderfolgenden Blöcken der Folge von Blöcken eine Informationsentität darstellt, unter Verwendung einer Folge von Fingerabdrücken für die Folge von Blöcken werden Identifikationsergebnisse für aufeinanderfolgende Fingerabdrücke bereitgestellt (12), wobei ein Identifikationsergebnis eine Zugehörigkeit eines Blocks von Informationseinheiten zu einer vorbestimmten Informationsentität darstellt. Hierauf werden wenigstens zwei Hypothesen aus den Identifikationsergebnissen für die aufeinanderfolgenden Fingerabdrücke gebildet (14), wobei eine erste Hypothese eine Annahme für die Zugehörigkeit der Folge von Blöcken zu einer ersten Informationsentität ist, und wobei die zweite Hypothese eine Annahme für die Zugehörigkeit der Folge von Blöcken zu der zweiten Informationsentität ist. Hierauf werden verschiedene Hypothesen untersucht (16), um ein

[Fortsetzung auf der nächsten Seite]

WO 2005/111998 A1



(81) **Bestimmungsstaaten** (soweit nicht anders angegeben, für jede verfügbare nationale Schutzrechtsart): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) **Bestimmungsstaaten** (soweit nicht anders angegeben, für jede verfügbare regionale Schutzrechtsart): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG,

ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Veröffentlicht:**

— mit internationalem Recherchenbericht

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

## Vorrichtung und Verfahren zum Analysieren eines Informationssignals

5

### Beschreibung

Die vorliegende Erfindung bezieht sich auf die Signalanalyse und insbesondere auf die Signalanalyse zu Zwecken der Identifikation eines Signalinhalts.

10

Um die immer größer werdenden Bestände an Audio- und Video-Material zu archivieren, leicht durchsuchbare Datenbanken zu erstellen oder über verschiedene Vertriebswege zu verteilen, werden automatische Informationserkennungssysteme benötigt, die helfen, Audio- und Videomaterial oder allgemein gesagt Informationsmaterial eindeutig, basierend auf dem Inhalt, zu identifizieren.

Eine Anwendung hierfür ist das sogenannte „Broadcast-Monitoring“ bzw. die Rundfunküberwachung. Mit Hilfe eines solchen Audio-Video-Überwachungssystems soll z. B. sichergestellt werden, daß nur rechtmäßige Inhalte verbreitet werden oder daß die jeweiligen Tantiemen für die Rechteinhaber des Audio- und Video-Materials korrekt abgerechnet werden.

25

Eine weitere Anwendung ist z. B. die Erkennung von Audiomaterial, welches über Peer-to-Peer-Netzwerke zwischen Partnern ausgetauscht werden soll.

30

Eine weitere Anwendung ist die Kontrollmöglichkeit der Werbeindustrie, um einen Fernseh- oder Radiosender dahingehend zu kontrollieren, ob die gebuchten Werbezeiten auch tat-

sächlich ausgestrahlt worden sind, oder ob nur Teile der gebuchten Werbeanteile ausgesendet worden sind, oder ob Teile der Werbesendungen während der Übertragung gestört worden sind, was z. B. in die Verantwortung des Fernseh- oder Rundfunk-Senders fallen könnte. An dieser Stelle sei 5 darauf hingewiesen, daß insbesondere die Kosten für die Fernsehwerbung in weit verbreiteten Programmen zu guten Sendezeiten derart hoch sind, daß die Werbeindustrie insbesondere im Hinblick auf diese hohen Kosten ein vitales Interesse an einer Kontrollmöglichkeit hat, um nicht nur auf 10 die Angaben der Rundfunksender vertrauen zu müssen. Bisher wird als Kontrollmöglichkeit auf bezahlte „Probehörer“ bzw. „Probehörerinnen“ oder „Probeseher“ bzw. „Probeseherinnen“ gesetzt, die ein bestimmtes Fernsehprogramm durchgehend betrachten und z. B. die genauen Zeiten protokollieren, zu 15 denen ein Werbespot übertragen wird und die ferner überwachen, ob während der Übertragung keine Störung aufgetreten ist, oder ob der gesamte Werbespot korrekt übertragen worden ist, ob also keine Bildverzerrung stattgefunden hat, 20 etc.

Die Nachteile dieses Konzepts sind evident. So sind zum einen die Kosten erheblich, und ist zum anderen die Verlässlichkeit oder Beweiskraft von Aussagen der Probehörer bzw. 25 Probeseher problematisch, insbesondere wenn Rückzahlungsforderungen in beträchtlicher Höhe erhoben werden, welche im Hinblick auf ihre Nachweisbarkeit allein an Probebeobachtern bzw. Probebeobachterinnen hängt.

30 Zum automatisierten Broadcast-Monitoring können verschiedene bekannte Systeme eingesetzt werden. So offenbart die WO 02/11123 A2 oder die Fachveröffentlichung: „Invited Talk: An Industrial-Strength Audio Search Algorithm“, Avery Wang,

ISMIR 2003, Baltimore, Oktober 2003, Systeme und Verfahren zum Erkennen von Ton- und Musiksignalen in einer Umgebung starken Rauschens und hoher Verzerrungen. Hierbei wird zunächst untersucht, ob eine Übereinstimmung zwischen Hash-  
5 Werten eines Referenz-Audio-Objekts und dem aktuell ermittelten Hash-Wert des noch unidentifizierten Audio-Objekts vorliegt. Ist dies der Fall, so wird der zugehörige Zeitoffset, also der relative Abstand vom Beginn des Audio-Objekts, des Hash-Werts in dem noch unidentifizierten Au-  
10 dio-Objekt und der Zeitoffset des Hash-Werts in dem Referenz-Audio-Objekt unter der jeweiligen Kennung des Referenz-Audio-Objekts gespeichert. Sind alle Eingangshashwerte verarbeitet, dann beginnt eine sogenannten Scanning-Phase. In dieser wird untersucht, wie viele Zeitoffsetpaare pro  
15 Referenz-Audio-Objektzeit kontinuierlich zusammenpassen. Wird eine bestimmte Anzahl festgestellt, so geht man von einer Identifizierung des entsprechenden Referenz-Audio-Objekts aus. Die Zeitoffsetpaare werden genau dann als zeitkontinuierlich, d. h. zeitlich zueinander gehörig, angesehen, wenn sie in einem zweidimensionalen Scatterplot mit einem Zeitoffset als x- und dem anderen als y-  
20 Koordinate eine Gerade bilden.

In der Fachveröffentlichung „Robust Audio Hashing for Content Identification“ von J. Haitzma, T. Kalker, J. Oostveen, in den Proceedings of the Content-Based Multimedia Indexing, 2001, [url:citeseer.ist.psu.edu/haitzma01robust.html](http://citeseer.ist.psu.edu/haitzma01robust.html), wird ein System zum robusten Audio-Hashing für eine Inhaltsidentifikation dargestellt. Zur  
25 inhaltsbasierten Musikerkennung wird eine Hash-Funktion eingesetzt, welche eine Bitfolge mit einem Abschnitt aus einem Audiosignal assoziiert, und zwar in einer Form, daß für die menschliche Schallwahrnehmung akustisch ähnliche  
30

Audiosignale auch eine ähnliche Bitfolge erzeugen. Zur Berechnung eines Hash-Werts wird zunächst das Audiosignal gefenstert und einer Transformation unterzogen, um schließlich eine Einteilung des Transformationsergebnisses in Frequenzbänder mit logarithmischer Bandbreite vorzunehmen. Für diese Frequenzbänder werden die Vorzeichen der Differenzen in Zeit- und Frequenzrichtung bestimmt. Die sich aus den Vorzeichen ergebende Bitfolge bildet den Hash-Wert. Es wird immer ein Hash-Wert für eine Audiosignallänge von 3 Sekunden berechnet. Liegt für einen solchen Abschnitt die Hamming-Distanz zwischen einem Referenz-Hash-Wert und einem zu untersuchenden Test-Hash-Wert unterhalb einer Schwelle  $s$ , so wird eine Übereinstimmung angenommen und der Testabschnitt dem Referenzelement zugeordnet.

15

Um eine Erkennung von Audiomaterial durchzuführen, wird das Audiosignal typischerweise in kleine Einheiten der Länge  $\Delta t$  zerlegt. Diese einzelnen Einheiten werden jeweils einzeln analysiert, um wenigstens eine gewisse zeitliche Auflösung zu haben.

20

Hieraus ergeben sich mehrere Probleme.

Die Erkennungsergebnisse der kleinen analysierten Zeitabschnitte des Audiosignals müssen so zusammengesetzt werden, daß für einen längeren Zeitabschnitt eine eindeutige, korrekte Aussage über das erkannte Audiosignal getroffen werden kann.

25

Zur Analyse eines kontinuierlichen Audiodatenstroms sollen Übergänge von einem Audioelement zu einem anderen richtig erkannt werden, d. h. ein Wechsel von einem Musikstück A zu einem Musikstück B.

30

Ferner existiert die Situation, bei der von einem Musikstück mehrere Versionen vorkommen, die z. B. gleich beginnen und sich erst nach einer gewissen Zeit anfangen zu unterscheiden. Hierbei sei beispielsweise an Kurzversionen oder Maxiversionen eines Songs gedacht. Alternativ existieren auch Situationen dahingehend, daß Musikstücke, die auf den gleichen Song zurückgehen, z. B. am Anfang unterschiedlich sind, einen identischen Mittelteil haben und sich gegen Ende wenigstens eines der beiden Musikstücke voneinander wieder unterscheiden. Für die Tantiemenvergabe an Urheberrecht-Inhaber kann es durchaus eine Rolle spielen, ob z. B. gegen eine höhere Gebühr die Maxiversion eines Songs gespielt werden kann, ob für eine mittlere Gebühr nur eine Normalversion gespielt werden darf, oder ob für eine geringe Gebühr bereits die Kurzversion eines Songs gespielt werden kann. In diesem Fall sollten verschiedene Versionen eines Songs zuverlässig voneinander unterschieden werden können.

Der vorstehende Stand der Technik ist dahingehend unzureichend, daß sich Erfassungsfehler ergeben, wenn die Ergebnisse der Einzelerkennungen einfach zusammengesetzt werden. Insbesondere werden keine Angaben dahingehend geliefert, ob und wie ein kontinuierlicher Audiodatenstrom aus mehreren verschiedenen Audioobjekten analysiert werden kann und wie entsprechende Wechsel zwischen verschiedenen Audioobjekten detektiert werden können. Des weiteren wird insbesondere im letztgenannten Stand der Technik die Vieldeutigkeit von Referenz-Hash-Werten zwar erwähnt. Es wird jedoch keine explizite Lösung für das Problem zur Bestimmung eines eindeutigen Kandidaten genannt. Gilt für einen Hash-Wert ein Audioobjekt als identifiziert, so wird für den unmittelbar

folgenden Hash-Wert lediglich geprüft, ob dieser zu dem identifizierten Audioobjekt paßt. Ist dies nicht der Fall, so wird neu gesucht, unter Einbeziehung von allen Referenz-audioobjekten.

5

Insbesondere zur Unterscheidung von unterschiedlichen Versionen ein und desselben Songs ist im Stand der Technik keine Lösung bekannt.

10 Die Aufgabe der vorliegenden Erfindung besteht darin, ein zuverlässiges Konzept zum Analysieren eines Informationssignals zu schaffen.

Diese Aufgabe wird durch eine Vorrichtung zum Analysieren  
15 eines Informationssignals gemäß Patentanspruch 1, ein Verfahren zum Analysieren eines Informationssignals gemäß Patentanspruch 20 oder ein Computer-Programm gemäß Patentanspruch 21 gelöst.

20 Der vorliegenden Erfindung liegt die Erkenntnis zugrunde, daß eine zuverlässige Inhaltsidentifikation dadurch erreicht wird, daß nicht nur Einzelerkennungsergebnisse für sich betrachtet berücksichtigt werden, sondern über einen gewissen Zeitraum hinweg. So existiert auch in der Folge  
25 von Einzelerkennungsergebnissen für eine Folge von Fingerabdrücken eine erhebliche Information, die zur Erkennung einsetzbar ist. Erfindungsgemäß wird daher aus einer Folge von Fingerabdrücken, die eine Folge von Blöcken eines Informationssignals darstellt, eine Bildung von wenigstens  
30 zwei unterschiedlichen Hypothesen durchgeführt, wobei eine erste Hypothese eine Annahme für die Zugehörigkeit der Folge von Blöcken zu einer ersten Informationsentität ist, und wobei die zweite Hypothese eine Annahme für die Zugehörig-

keit der Folge von Blöcken zu der zweiten Informationsentität ist. Die wenigstens zwei Hypothesen werden nunmehr untersucht und einer Bewertung dahingehend unterzogen, daß eine Aussage über das Informationssignal basierend auf einem Untersuchungsergebnis getroffen wird. Die Aussage könnte beispielsweise darin bestehen, daß bestimmt wird, daß die Folge von Blöcken eine Informationsentität darstellt, die eine Hypothese hat, die am wahrscheinlichsten ist. Die Aussage könnte alternativ oder zusätzlich darin bestehen, daß eine Informationseinheit mit dem Fingerabdruck beendet ist, der als zeitlich letzter Fingerabdruck aus der Folge von Fingerabdrücken zu der wahrscheinlichsten Hypothese beiträgt.

Vorzugsweise werden die Hypothesen dahingehend untersucht, daß es für Fingerabdrücke wenigstens zwei unterschiedliche Identifikationsergebnisse gibt, sowie daß es für jeden der zwei unterschiedlichen Identifikationsergebnisse ein Zuverlässigkeitsmaß gibt, wobei dieses Zuverlässigkeitsmaß in einer konkreten Zahl bestehen kann. Dieses Zuverlässigkeitsmaß kann jedoch auch implizit dahingehend gegeben sein, daß allein die Tatsache, daß z. B. zwei Identifikationsergebnisse geliefert werden, eine Zuverlässigkeit von z. B.  $1/2$  signalisiert wird und daß diese Zahl nicht explizit angegeben wird.

Zur Beurteilung, ob eine Hypothese wahrscheinlicher ist als die andere Hypothese werden vorzugsweise Zuverlässigkeitsmaße der Einzelerkennungen für die betreffende Anzahl von zeitlich aufeinanderfolgenden Blöcken zusammengefaßt, wobei diese Zusammenfassung vorzugsweise in einer Aufaddition bestehen wird. Hierauf wird die Hypothese, die das höchste

zusammengefaßte Zuverlässigkeitsmaß liefert, als die wahrscheinlichste Hypothese bewertet.

Bei einem bevorzugten Ausführungsbeispiel der vorliegenden  
5 Erfindung wird als Mittel zum Bereitstellen von aufeinanderfolgenden Identifikationsergebnissen eine Fingerabdruck-Datenbank verwendet, in der eine Anzahl von Referenz-Fingerabdrücken jeweils in Zuordnung zu einem Identifikationsergebnis abgelegt sind. Hierauf wird mit dem aus einem  
10 Block des zu analysierenden Informationssignals erzeugten Fingerabdrucks eine Datenbanksuche vorgenommen, um innerhalb der Datenbank nach einem Referenzfingerabdruck zu suchen, der eine Übereinstimmung mit dem Test-Fingerabdruck liefert. Je nach Ausgestaltung der Datenbank wird als  
15 Suchergebnis nur der beste Treffer, also der Treffer mit einem minimalen Abstandsmaß von der Datenbank als Identifikationsergebnis ausgegeben. Es werden auch Datenbanken bevorzugt, die nicht nur qualitativ ein Trefferergebnis liefern, sondern auch ein quantitatives  
20 Trefferergebnis liefern, dahingehend, daß eine Anzahl möglicher Treffer mit einem zugeordneten Zuverlässigkeitsmaß ausgegeben wird, so daß beispielsweise alle Treffer mit einem Zuverlässigkeitsmaß, das höher oder gleich einer bestimmten Schwelle, wie beispielsweise 20%  
25 ist, von der Datenbank ausgegeben werden.

Bei dem bevorzugten Ausführungsbeispiel der vorliegenden Erfindung wird dann eine neue Hypothese begonnen, wenn ein neues Identifikationsergebnis auftaucht, für das noch keine Hypothese existiert. Dieses Prozedere wird für eine bestimmte Anzahl von Blöcken durchgeführt, um dann in die  
30 Vergangenheit gerichtet zu untersuchen, ob eine bestimmte sich als zuverlässig herausgestellte Hypothese bereits zu

Ende ist, um dann diese Hypothese als die wahrscheinlichste Hypothese zu erkennen.

Ein Vorteil der vorliegenden Erfindung besteht darin, daß  
5 das Konzept zuverlässig arbeitet und dennoch fehlertolerant insbesondere im Hinblick auf Übertragungsfehler ist. So wird nicht anhand eines einzigen Blockes versucht eine Entscheidung herbeizuführen, sondern es wird eine Folge von aufeinanderfolgenden Blöcken durch Hypothesenbildung gewissermaßen gemeinsam betrachtet und ausgewertet, so daß nicht  
10 kurzfristige Übertragungsstörungen bzw. allgemeines auftretendes Rauschen den gesamten Erkennungsprozeß hinfällig machen.

15 Weiterhin liefert das erfindungsgemäße Konzept automatisch eine Protokollierung der Übertragungsqualität vom Anfang bis zum Ende z. B. eines Werbespots. Auch wenn eine Hypothese als wahrscheinlichste Hypothese erkannt worden ist, wenn also bestimmt wird, daß ein bestimmter Werbespot da  
20 war, so sind dennoch Qualitätsschwankungen innerhalb des Werbespots anhand der Zuverlässigkeitsmaße nachvollziehbar. Ferner ist damit insbesondere die komplette zeitliche Kontinuität eines Werbespots als Beispiel für eine Informationsentität nachvollziehbar und protokollierbar, und zwar  
25 insbesondere dahingehend, daß nicht irgend etwas aus dem Werbespot ständig wiederholt worden ist, sondern daß der gesamte Werbespot vom Anfang des Werbespots bis zum Ende des Werbespots kontinuierlich gesendet worden ist.

30 Die vorliegende Erfindung ist ferner dahingehend vorteilhaft, daß durch Hypothesenbildung automatisch das Ende einer Informationsentität und der Beginn einer Informationsentität erkannt werden. Dies liegt daran, daß eine Zuord-

nung zu einer Informationsentität im allgemeinen eindeutig sein wird. Dies bedeutet, daß über einen bestimmten Zeitpunkt nicht mehrere Informationsentitäten gemeinsam abgespielt werden können, sondern daß zumindest für die über-  
5 große Anzahl von Programminhalten nur eine Informationsentität im Informationssignal zu einem Zeitpunkt enthalten ist. Die Hypothesenuntersuchung und die Beurteilung der Hypothesen aufgrund der Hypothesenuntersuchung liefert automatisch einen Zeitpunkt, zu dem eine vorherige Informati-  
10 onsentität beendet ist und zu dem eine neue Informationsentität beginnt. Dies liegt an der bis in die Hypothesen hinein beibehaltenen Blockzuordnung. So entspricht nach wie vor eine Folge von Fingerabdrücken einer Folge von Blöcken, und entspricht wiederum eine Folge von Identifikationsergebnissen einer Folge von Fingerabdrücken, so daß eine  
15 Hypothese dem ursprünglichen Informationssignal in zeitlicher Hinsicht eindeutig zugeordnet ist.

Das erfindungsgemäße Konzept ist ferner dahingehend vor-  
20 teilhaft, daß keine „Unentschieden“-Situationen zwischen zwei Hypothesen entstehen, auch wenn Informationsentitäten streckenweise dennoch das identische Audiomaterial haben, wie beispielsweise bei Kurzversionen oder Langversionen ein und desselben Songs.

25

Bevorzugte Ausführungsbeispiele der vorliegenden Erfindung werden nachfolgend Bezug nehmend auf die beiliegenden Zeichnungen detailliert erläutert. Es zeigen:

30 Fig. 1 ein Blockschaltbild einer erfindungsgemäßen Vorrichtung;

- Fig. 2 ein Blockschaltbild einer für das in Fig. 1 gezeigte Ausführungsbeispiel verwendbaren Datenbank;
- 5 Fig. 3 eine schematische Darstellung eines Ausgabeergebnisses für eine Folge von Fingerabdrücken für eine Folge von Zeitintervallen sowie die zugeordneten Hypothesen;
- 10 Fig. 4a-4c ein Beispielszenario für nachfolgende Anwendungsbeispiele;
- Fig. 5a-5d eine schematische Darstellung verschiedener Fehltauswertungen;
- 15 Fig. 6 ein Blockschaltbild eines bevorzugten Ausführungsbeispiels der vorliegenden Erfindung;
- Fig. 7a-7c eine Darstellung der Funktionalität des erfindungsgemäßen Konzepts für das in den Fig. 4a-4c dargestellte Ausgangsszenario;
- 20 Fig. 8 eine schematische Darstellung eines Informationssignals mit Informationseinheiten, Blöcken von Informationseinheiten und Informationseinstitäten mit einer Mehrzahl von Blöcken;
- 25 Fig. 9 ein bekanntes Szenario zum Aufbauen einer Fingerabdruck-Datenbank; und
- 30 Fig. 10 ein bekanntes Szenario zum Audio-Identifizieren mittels einer gemäß Fig. 9 geladenen Fingerabdruck-Datenbank.

Fig. 1 zeigt ein Blockschaltbild einer Vorrichtung zum Analysieren eines Informationssignals gemäß einem bevorzugten Ausführungsbeispiel der vorliegenden Erfindung. Ein beispielhaftes Informationssignal ist bei 800 in Fig. 8 dargestellt. Das Informationssignal 800 besteht aus einer Folge 802 von zeitlich aufeinanderfolgenden Blöcken von Informationseinheiten, wobei die einzelnen Informationseinheiten 804 z. B. Audio-Abtastwerte, Video-Pixels oder Video-Transformationskoeffizienten etc. sein könnten. Eine Mehrzahl von Blöcken der Folge 802 bilden zusammen immer eine Informationsentität 806. Bei dem in Fig. 8 gezeigten Ausführungsbeispiel bilden die ersten sechs Blöcke die erste Informationsentität, und bilden die Blöcke 7, 8, 9, 10 die zweite Informationsentität. Ab den Blöcken 11 bis n ist in Fig. 8 z. B. eine dritte Informationsentität dargestellt. Eine Informationsentität könnte beispielsweise ein Musikstück, eine gesprochene Passage, ein Videobild oder auch z. B. ein Teil eines Videobildes sein. Eine Informationsentität könnte jedoch auch ein Text oder z. B. eine Seite eines Textes sein, wenn das Informationssignal auch Textdaten umfaßt.

Die in Fig. 1 gezeigte Vorrichtung ist ausgebildet, um unter Verwendung einer Folge von Fingerabdrücken FA1, FA2, FA3, ..., FAi zu arbeiten, die aus der Folge von Blöcken 802 erzeugt werden, oder die z. B. von einem Speicher abgerufen werden, wenn die Fingerabdrücke bereits vor der Analyse erzeugt worden sind oder vielleicht sogar mit dem Informationssignal mitgeliefert werden, je nach Ausführungsform. Es sei darauf hingewiesen, dass bei der Blockbildung auch Blocküberlappungstechniken eingesetzt werden können, wie sie z. B. aus der Audiocodierung bekannt sind.

Auf jeden Fall arbeitet die Vorrichtung zum Analysieren des Informationssignals unter Verwendung einer Folge von Fingerabdrücken für die Folge von Blöcken, so daß die Folge von Blöcken 802 durch die Folge von Fingerabdrücken FA1, FA2, FA3, FA4, ..., FA<sub>i</sub> dargestellt ist. Die Folge von Fingerabdrücken wird in einen Fingerabdruck-Eingang in eine Einrichtung 12 zum Bereitstellen von Identifikationsergebnissen für aufeinanderfolgende Fingerabdrücke eingespeist.

Die Einrichtung 12 zum Bereitstellen von aufeinanderfolgenden Identifikationsergebnissen ist wirksam, um für die aufeinanderfolgenden Fingerabdrücke aufeinanderfolgende Identifikationsergebnisse zu liefern, wobei ein Identifikationsergebnis eine Zugehörigkeit eines Blocks von Informationseinheiten zu einer vorbestimmten Informationsentität darstellt. Wenn z. B. davon ausgegangen wird, daß ein Song eine zeitliche Länge hat, die in etwa sechs Blöcken entspricht, so liefern die sechs Blöcke zwar unterschiedliche Fingerabdrücke, in der Einrichtung 12 zum Bereitstellen ist jedoch signalisiert, daß alle diese sechs Blöcke zu der vorbestimmten Informationsentität, also zu dem besagten Lied gehören.

Je nach Ausführungsform wird die Einrichtung 12 zum Bereitstellen für einen Fingerabdruck einen oder mehrere Identifikationsergebnisse liefern. Dieser eine oder die mehreren Identifikationsergebnisse werden einer Einrichtung 14 zum Bilden von wenigstens zwei Hypothesen aus den Identifikationsergebnissen für die aufeinanderfolgenden Fingerabdrücke zugeführt. Insbesondere stellt eine erste Hypothese eine Annahme für die Zugehörigkeit der Folge von Blöcken zu einer ersten Informationsentität dar, und ist die zweite Hypothese eine Annahme für die Zugehörigkeit der Folge von

Blöcken zu der zweiten Informationsentität. Die verschiedenen Hypothesen H1, H2, ... werden einer Einrichtung 16 zum Untersuchen der Hypothesen zugeführt, wobei die Einrichtung 16 ausgebildet ist, um gemäß eines einstellbaren Untersuchungsalgorithmus zu arbeiten, um schließlich ein Untersuchungsergebnis an einem Untersuchungsergebnisausgang 18 zu liefern.

Dieses Untersuchungsergebnis an der Leitung 18 wird dann einer Einrichtung 20 zum Treffen einer Aussage über das Informationssignal zugeführt. Die Einrichtung 20 zum Treffen einer Aussage über das Informationssignal ist ausgebildet, um eine Information über das Informationssignal basierend auf dem Untersuchungsergebnis auszugeben und kann vielerlei Einstellungen haben.

Für alle Einstellungen ist gemeinsam, daß die Aussage über das Informationssignal auf der Basis des Untersuchungsergebnisses 18 getroffen wird. Beispiele für verschiedene Aussagen über das Informationssignal bestehen darin, daß bestimmt wird, daß die Folge von Blöcken eine Informationsentität darstellt, die eine Hypothese hat, die am wahrscheinlichsten ist. Alternative Aussagen bestehen darin, daß eine Informationsentität mit dem Fingerabdruck beendet ist, der als zeitlich letzter Fingerabdruck zu der wahrscheinlichsten Hypothese beiträgt. Eine alternative Aussage, die durch die Einrichtung 20 getroffen werden kann, besteht darin, daß festgestellt wird, daß in dem Informationssignal eine Informationsentität an sich vorhanden ist oder nicht.

Das erfindungsgemäße Postprocessing, das insbesondere durch die Einrichtungen 14, 16 und 20 geliefert wird, also das

Bilden von wenigstens zwei Hypothesen, das Untersuchen der Hypothesen und das Treffen einer Aussage auf der Basis eines Untersuchungsergebnisses ermöglicht somit nicht nur die Identifikation eines Stücks in einem an sich unbekanntem, also zu analysierenden Informationssignal, sondern ermöglicht auch - abgesehen von der Identifikation eines Stücks an sich - die Erkennung des Endes eines ersten Stücks, also einer ersten Informationsentität, und das Detektieren des Beginns einer zweiten, auf die erste Informationsentität folgenden Informationsentität.

Im Hinblick auf die Werbeüberwachung liefert das erfindungsgemäße Postprocessing-Konzept jedoch auch die Möglichkeit, zu erfassen, ob ein bestimmtes Stück in dem Informationssignal vorhanden war oder nicht. Hierbei würden die aus dem Informationssignal gewonnenen Fingerabdrücke lediglich mit einem Satz von Fingerabdrücken, nämlich dem Satz von Fingerabdrücken, die die vorbestimmte Informationsentität, also einen bestimmten Werbespot darstellen, verglichen. Diese Aussage ist somit nicht primär dahingehend zu sehen, daß eine Informationsentität identifiziert wird, oder daß das Ende einer Informationsentität und der Beginn einer folgenden Informationsentität erfaßt werden, sondern besteht darin, zu erfassen, ob eine bestimmte Informationsentität in einem an sich unbekanntem zu analysierenden Informationssignal vorhanden ist oder nicht.

Fig. 2 zeigt eine spezielle bevorzugte Ausführungsform der Einrichtung 12 zum Bereitstellen von Identifikationsergebnissen für aufeinanderfolgende Fingerabdrücke. Bei einem bevorzugten Ausführungsbeispiel umfaßt die Einrichtung 12, wie es in Fig. 2 gezeigt ist, eine Datenbank, die verschiedene Referenzfingerabdrücke FAR<sub>j</sub> umfaßt, die alle in Zuord-

nung zu einem Identifikationsergebnis, d. h. ID<sub>k</sub>, abgespei-  
chert sind. Bei dem bevorzugten Ausführungsbeispiel werden  
die Fingerabdrücke FA<sub>i</sub> nach und nach, also zeitlich sequen-  
tiell verarbeitet. So wird ein Fingerabdruck FA<sub>i</sub> in die Da-  
5 tenbank über eine Eingangsleitung 24 eingespeichert. In der  
Datenbank wird dann der eingespeicherte Fingerabdruck FA<sub>i</sub>  
mit allen Referenzfingerabdrücken FA<sub>rj</sub> verglichen. bei dem  
bevorzugten Ausführungsbeispiel ist die Datenbank nicht ei-  
ne qualitative Datenbank, die feststellt, daß ein Eingangs-  
10 fingerabdruck mit einem gespeicherten Referenzfingerabdruck  
übereinstimmt oder nicht, sondern ist die Datenbank eine  
quantitative Datenbank, die ein Abstandsmaß bzw. ein Zuver-  
lässigkeitsmaß für die ausgegebenen Ergebnisse liefern  
kann. So würde die Datenbank 22 bei dem in Fig. 2 gezeigten  
15 bevorzugten Ausführungsbeispiel an ihrem Ausgang 26 z. B.  
das Ergebnis liefern, das in einer Ergebnistabelle 28 dar-  
gestellt ist. So würde die Datenbank z. B. sagen, daß der  
Fingerabdruck FA<sub>i</sub> auf ein Identifikationsergebnis ID<sub>x</sub>, also  
auf ein Musikstück beispielsweise x mit einer Zuverlässig-  
20 keit ZV<sub>1</sub> von 60% hinweist. Gleichzeitig wird die Datenbank  
jedoch auch sagen, daß der Fingerabdruck FA<sub>i</sub> mit einer Zu-  
verlässigkeit von 50% auf ein Stück mit dem Identifikati-  
onsergebnis ID<sub>y</sub> hinweist. Schließlich könnte die Datenbank  
auch ausgeben, daß der Fingerabdruck FA<sub>i</sub> mit einem Zuver-  
25 lässigkeitsmaß ZV<sub>3</sub> von beispielsweise 40% auf ein wieder  
anderes Stück mit der Identifikation ID<sub>z</sub> hinweist.

Je nach Ausführungsform kann die gesamte Ergebnistabelle 28  
der Einrichtung 14 zum Bilden von wenigstens zwei Hypothe-  
30 sen von Fig. 1 zugeführt werden. Alternativ könnte jedoch  
die Datenbank 22 selbst bereits eine Entscheidung treffen  
und immer nur den wahrscheinlichsten Wert, also im vorlie-  
genden Fall das Ergebnis ID<sub>x</sub> der Einrichtung 14 zum Bilden

von wenigsten zwei Hypothesen zuführen. In diesem Fall müßte nicht unbedingt das Zuverlässigkeitsmaß ZV1 ebenfalls der Einrichtung 14 zum Bilden von wenigstens zwei Hypothesen zugeführt werden. Vielmehr könnte auf die weitere Übermittlung der Zuverlässigkeitsmaße ZVi verzichtet werden. Alternativ könnte jedoch auch die Einrichtung 12 zum Bereitstellen der Identifikationsergebnisse, die gleichzeitig auch die Zuverlässigkeitsmaße liefert, ausgebildet sein, um die Zuverlässigkeitsmaße ZVi in entsprechender Reihenfolge in Zuordnung zu den Blöcken nicht der Einrichtung 14 zum Bilden von wenigstens zwei Hypothesen zuzuführen, sondern der Einrichtung 16 zum Untersuchen der Hypothesen, da diese Einrichtung 16 die Zuverlässigkeitsmaße erst benötigt, um z. B. die wahrscheinlichste Hypothese zu finden.

15

Aus der Datenbank 22 in Fig. 2 ist zu sehen, daß ein Identifikationsergebnis, wie z. B. ID1 mehrere zugeordnete Fingerabdrücke FAR11, FAR12, FAR13 haben kann, was darauf hindeutet, daß das Stück, das durch ID1 identifiziert wird, mehrere Blöcke hat. Je nach Implementierung kann jedoch auch ein einziger langer Fingerabdruck für das Stück mit der Identifikation ID1 gespeichert sein, der sich jedoch aus den einzelnen Fingerabdrücken FAR11, FAR12, FAR13, .... zusammensetzt. Die Datenbank würde dann den eingespeisten Fingerabdruck FAi, der von der Blocklänge abhängt und typischerweise viel kürzer ist als der lange Fingerabdruck, mit dem langen Fingerabdruck in jeder Zeile der Datenbank korrelieren, um festzustellen, ob ein Ausschnitt des langen abgespeicherten Referenzfingerabdrucks mit dem an der Leitung 24 eingespeisten Referenzfingerabdruck FAi übereinstimmt oder nicht. Das Zuverlässigkeitsmaß würde sich hierbei gewissermaßen automatisch ergeben, nämlich einfach durch quantitative Auswertung des Korrelationsergebnisses.

Ferner sei Bezug nehmend auf Fig. 2 bereits auf die beiden letzten Zeilen verwiesen, die mit den Identifikationsergebnissen ID108 und ID109 bezeichnet sind. ID108 bezeichnet  
5 eine lange Version eines Musikstücks, wie es noch Bezug nehmend auf Fig. 4a erläutert wird, während ID109 eine kurze Version desselben Musikstücks identifiziert, wie es in Fig. 4b dargestellt ist.

10 Wie es bereits ausgeführt worden ist, kann die Datenbank 22, also diese Implementierung der Einrichtung 12 zum Bereitstellen von Identifikationsergebnissen für aufeinanderfolgende Fingerabdrücke derart ausgebildet sein, daß sie immer nur das wahrscheinlichste Identifikationsergebnis  
15 liefert. Alternativ könnte die Datenbank 22 jedoch auch ausgebildet sein, um z. B. immer nur die Identifikationsergebnisse zu liefern, deren Wahrscheinlichkeit höher als eine Mindestschwelle ist, wie beispielsweise eine Schwelle von 5%. Dies würde dazu führen, daß die Anzahl der Zeilen  
20 der Tabelle von Fingerabdruck zu Fingerabdruck variieren wird. Wieder alternativ könnte die Datenbank 22 jedoch auch implementiert sein, um für jeden Eingangsfingerabdruck FAi eine bestimmte Anzahl von wahrscheinlichsten Kandidaten, wie beispielsweise die „Top Ten“, also die zehn wahrscheinlichsten  
25 Kandidaten der Einrichtung 14 zum Bilden von wenigsten zwei Hypothesen zuzuführen.

Nachfolgend wird anhand von Fig. 3 eine Ausführungsform der Datenbank 22 dargestellt, bei der die Datenbank immer die  
30 drei wahrscheinlichsten Identifikationsergebnisse samt zugehörigerer Zuverlässigkeitswerte der Einrichtung 14 zum Bilden von Hypothesen zuführt, also gewissermaßen eine „Top Three“-Implementierung umfaßt. So ist aus Fig. 3 zu sehen,

daß für den Fingerabdruck FA1 Identifikationsergebnisse ID1, ID2, ID3 geliefert werden, und zwar mit den jeweiligen Zuverlässigkeitsmaßen 40%, 60% oder 30%. Für das Zeitintervall  $\Delta t_2$ , also für den Fingerabdruck FA2 wird wieder eine  
5 Lieferung der Identifikationsergebnisse ID1, ID2, ID3 stattfinden, nun jedoch mit einer anderen jeweiligen Wahrscheinlichkeit, also mit einem anderen jeweiligen Zuverlässigkeitsmaß, das lediglich beispielhaft in Fig. 3 prozentual dargestellt ist. Dieses Prozedere wird für sämtliche  
10 Eingangs-Fingerabdrücke FA1 bis FA8 durchgeführt. Der Einrichtung 14 zum Bilden von wenigstens zwei Hypothesen, wie es in Fig. 1 dargestellt ist, werden diese Identifikationsergebnisse geliefert. Die Einrichtung 14 zum Bilden von wenigstens zwei Hypothesen ist ausgebildet, um immer dann,  
15 wenn ein neues Identifikationsergebnis von der Einrichtung 12 zum Bereitstellen der Identifikationsergebnisse geliefert wird, eine neue Hypothese zu beginnen. Dies ist aus Fig. 3 daraus ersichtlich, daß beim Zeitpunkt  $\Delta t_1$  die Hypothesen H1, H2, H3 mit ID1, ID2 bzw. ID3 begonnen werden,  
20 und daß im Zeitintervall  $\Delta t_7$  wieder neue Hypothesen mit ID108, ID109, ID4 begonnen werden, und daß im Zeitintervall  $\Delta t_8$  aufgrund der Tatsache, daß dort bei dem gezeigten Beispiel ID8 zum ersten Mal auftritt, eine weitere Hypothese H4 für ID8 begonnen wird.

25

Die Einrichtung 14 zum Bilden von wenigstens zwei Hypothesen ist daher wirksam, um für jeden neuen Fingerabdruck zu sehen, ob ein neues Identifikationsergebnis kommt, um eine neue Hypothese zu beginnen, und um eine bereits früher begonnene Hypothese dahingehend weiterzuführen, wenn für einen  
30 Zeitabschnitt  $\Delta t_i$  für die bereits früher begonnene Hypothese in den „Top Three“ oder „Top x“ ein Element enthalten ist, das, wenn auch mit geringerer Wahrscheinlich-

keit dennoch ein Identifikationsergebnis für eine gerade  
begonnene Hypothese liefert. Dieses Prozedere wird eine be-  
stimmte Zeit lang fortgeführt. Dann, beispielsweise zu vor-  
bestimmten Zeitpunkten oder, getriggert durch einen Benut-  
5 zer, etc. wird die Einrichtung 16 zum Untersuchen der Hypo-  
thesen die für die Vergangenheit gebildeten Hypothesen un-  
tersuchen und für den in Fig. 3 gezeigten Fall beispiels-  
weise die Zuverlässigkeitsmaße der Hypothesen H1, H2, H3  
für die Zeiträume  $\Delta t_1$  bis  $\Delta t_6$  addieren. Die Einrichtung 16  
10 zum Untersuchen von wenigstens zwei Hypothesen würde dann  
feststellen, daß das Stück am wahrscheinlichsten ID1 ist,  
daß also die Hypothese H1 für den Zeitraum  $\Delta t_1$  bis  $\Delta t_6$  die  
wahrscheinlichste Hypothese ist, da das Zuverlässigkeitsmaß  
einen Wert von 420 erreicht, während für die zweite Hypo-  
15 these lediglich ein Zuverlässigkeitsmaß von 230 erreicht  
wird, und während für die dritte Hypothese lediglich ein  
Zuverlässigkeitsmaß von 135 erreicht wird.

Bei dem in Fig. 3 gezeigten Fall starten alle drei Hypothe-  
20 sen gleichzeitig und enden alle drei Hypothesen gleichzei-  
tig. Dies muß jedoch nicht zwingend so sein. So könnte die  
Hypothese H1 z. B. früher enden, also z. B. zum Zeitpunkt  
 $\Delta t_5$ . Das Zuverlässigkeitsmaß von ID1 müßte in diesem Fall  
um 90 reduziert werden, womit man bei einem Wert von 330  
25 ankommen würde. In diesem Fall würde sich ergeben, daß die  
Hypothese H1 trotzdem die wahrscheinlichste Hypothese ist,  
obgleich Hypothese H2 über einen längeren Zeitraum, jedoch  
insgesamt mit geringerer Wahrscheinlichkeit vorhanden ist.  
Aus dem in Fig. 3 gezeigten Beispiel ist ferner zu sehen,  
30 daß die Hypothese H1 trotz der Tatsache, daß sie für  $\Delta t_1$   
weniger wahrscheinlich war als die Hypothese H2, am Ende  
„die Nase vorn“ hat.

Aus Fig. 3 ist ferner auch zu sehen, daß eine Hypothese auch „Löcher“ dahingehend haben könnte, daß beispielsweise im Zeitintervall  $\Delta t_4$  aus irgendwelchen Gründen, beispielsweise aus Gründen der Störung eines Übertragungskanals etc. lediglich ID2 und ID3 mit vernünftiger Wahrscheinlichkeit geliefert werden, nicht jedoch ID1. Dann müßte der Zuverlässigkeitswert für ID1 um 60 reduziert werden, was wiederum dahingehend zum Ergebnis haben würde, daß die Gesamtzuverlässigkeit statt 420 360 wäre, so daß auch in diesem Fall die Hypothese H1 die wahrscheinlichste Hypothese ist.

Aus den vorstehend beschriebenen Szenarien ist daher zu sehen, daß das erfindungsgemäße Konzept, das auf der Basis eines Postprocessings mit Hypothesen arbeitet und einerseits die Sequenz und andererseits die Zuverlässigkeitsmaße der einzelnen Fingerabdruckidentifikationsvorgänge berücksichtigt, außerordentlich robust gegenüber Übertragungsfehlern und auch gegenüber problematischen Funktionalitäten in der Datenbank oder auch gegenüber Fingerabdrücken ist, die sich für manche Informationsentitäten, wie z. B. Musikstücke, Videobilder, Texte, etc. vielleicht nicht so stark voneinander unterscheiden, wie man es sich wünschen würde.

Bei einem bevorzugten Ausführungsbeispiel ist eine Hypothese ein gespeichertes Protokoll (Fig. 3: H1, H2, H3, ...), vorzugsweise in der Form einer gespeicherten Liste, das einerseits einen Hinweis auf die Informationsentität, für die die Hypothese aufgestellt wird, und andererseits einen Hinweis auf Fingerabdrücke bzw. Blöcke von Informationseinheiten, für die die Hypothese geführt wird, aufweist. Vorzugsweise enthält das Protokoll auch noch für einen Block bzw. Fingerabdruck ein Zuverlässigkeitsmaß.

Aus Fig. 3 ist ferner zu erkennen, daß sich die erste Informationsentität nur über den Zeitraum  $\Delta t_1$  bis  $\Delta t_6$  erstreckt und ab  $\Delta t_7$  eine neue Entität beginnt. Dies ist insbesondere auch daraus ersichtlich, daß alle drei Hypothesen  
5 gleichzeitig enden, bzw. daß, auch wenn die Hypothese H3 beispielsweise sich noch über  $\Delta t_7$  erstreckt hätte, nunmehr völlig andere Identifikationswerte mit sehr hoher Wahrscheinlichkeit, nämlich ID108 und ID109 mit Wahrscheinlichkeiten von 90 und 85 auftreten und damit die „klaren Gewinner“ aus dem vorherigen Zeitraum „ablösen“.  
10

Am Ende von Fig. 3 sind die verschiedenen Aussagen, die beispielhaft getroffen werden können, dargestellt, nämlich daß die Informationsentität im Zeitraum  $\Delta t_1$  bis  $\Delta t_6$  das  
15 durch ID1 identifizierte Musikstück ist. Alternativ könnte die Aussage auch dahingehend lauten, daß ein Informationsentität-Wechsel zwischen  $\Delta t_6$  und  $\Delta t_7$  stattfindet. Alternativ könnte eine Aussage jedoch auch dahingehend lauten, daß das Musikstück, das durch ID1 identifiziert ist, im Informationssignal enthalten ist.  
20

Nachfolgend wird Bezug nehmend auf die Fig. 9 und 10 zunächst allgemeiner auf Datenbanksysteme eingegangen, wie sie in Verbindung mit der vorliegenden Erfindung vorteilhaft eingesetzt werden können. Die vorliegende Erfindung  
25 basiert somit auf einem System zu Identifikation von Audiomaterial, wie beispielsweise Musik. Das System kennt zwei Betriebsphasen. In der Trainings-Phase, die anhand von Fig. 9 dargestellt ist, lernt das Erkennungssystem die später zu identifizierenden Stücke. In der Identifikations-Phase, die  
30 in Fig. 10 dargestellt ist, können die vorher trainierten Audiostücke wiedererkannt werden.

Um ein Stück Musik - oder auch jedes andere Tonsignal - zu identifizieren, wird aus diesem ein kompakter und einzigartiger Datensatz extrahiert, der auch als Fingerprint oder Signatur bezeichnet wird. Diese Extraktion findet in einer Block Merkmalsextraktion 900 statt. In der Trainings- oder Einlernphase werden von einem Satz bekannter Audioobjekte solche Fingerprints erstellt und in einer Fingerprint-Datenbank 902 abgelegt. Vorzugsweise ist die Merkmalsextraktionseinrichtung 900 ausgebildet, um als Merkmal das SFM-Merkmal zu verwenden, wobei SFM für „Spectral Flatness Measure“ steht. Selbstverständlich können auch andere Fingerabdruckserzeugungssysteme bzw. Merkmalsextraktionsergebnisse eingesetzt werden. Es hat sich jedoch herausgestellt, daß Tonalitäts-bezogene Merkmale und insbesondere das SFM-Merkmal einerseits eine besonders gute Unterscheidungskraft und andererseits eine besonders gute Kompaktheit haben. Zu diesem Zweck wird jeder Block zunächst einer Zeit-Frequenz-Umsetzung unterzogen, um dann mit den aus der Zeit-Frequenz-Umsetzung erzeugten Werten gemäß folgender Gleichung ein SFM für einen Block zu berechnen.

$$SFM = \frac{\left[ \prod_{n=0}^{N-1} X(n) \right]^{\frac{1}{N}}}{\frac{1}{N} \sum_{n=0}^{N-1} X(n)}$$

In dieser Gleichung stellt  $X(n)$  das Quadrat eines Absolutwerts einer Spektralkomponente mit den Index  $n$  dar, wobei  $N$  die Gesamtanzahl der Spektralkoeffizienten eines Spektrums darstellt. Es ist aus der Gleichung zu sehen, daß das SFM-Maß gleich dem Quotienten aus dem geometrischen Mittelwert der Spektralkomponenten und dem arithmetischen Mittelwert der Spektralkomponenten ist. Es ist bekannt, daß der geo-

metrische Mittelwert immer kleiner oder höchstens gleich dem arithmetischen Mittelwert ist, so daß das SFM einen Wertebereich zwischen 0 und 1 hat. In diesem Kontext zeigt ein Wert nahe 0 ein tonales Signal an, und zeigt ein Wert  
5 nahe 1 ein ziemlich rauschartiges Signal mit einer flachen Spektralkurve an. Es sei darauf hingewiesen, daß der arithmetische Mittelwert und der geometrische Mittelwert nur gleich sind, wenn alle  $X(n)$  identisch sind, was einem vollständig atonalen, d. h. rauschartigen oder pulsartigen Signal  
10 entspricht. Wenn in einem Extremfall jedoch nur eine Spektralkomponente einen sehr hohen Wert hat, während andere Spektralkomponenten  $X(n)$  sehr kleine Werte haben, wird das SFM-Maß einen Wert nahe 0 haben, was ein sehr tonales Signal andeutet.

15

Das SFM-Konzept ist sowie andere Merkmalsextraktionskonzepte, um Fingerabdrücke zu erzeugen, sind z. B. in der WO 03/007185 dargestellt.

20 In der Identifikations-Phase, die in Fig. 10 dargestellt ist, findet ebenfalls typischerweise dieselbe Merkmalsextraktion 900 wie bei der Trainings-Phase statt. Insbesondere wird der Fingerprint, welcher von dem Audio-Objekt am Audioeingang für einen Zeitabschnitt  $\Delta t$  extrahiert wurde,  
25 mit den Referenz-Fingerprints der Fingerprint-Datenbank 902 mittels eines Vergleichers 904 verglichen, wobei der Vergleich typischerweise in der Einrichtung 12 zum Bereitstellen von Identifikationsergebnissen, wie sie anhand von Fig. 1 dargestellt worden ist, enthalten sein wird.  
30 Hierauf erhält man bei Feststellung einer Übereinstimmung anhand eines bestimmten Kriteriums ein Erkennungsergebnis für den Zeitabschnitt  $\Delta t$ . Wird also eine Übereinstimmung anhand eines bestimmten Kriteriums festgestellt, so kann

der unbekannte Fingerprint, und damit der Ausschnitt aus dem unbekanntem Audio-Objekt, einem Referenzmaterial in der Datenbank, also einer Liste von Identifikationsergebnissen  $ID_i$ ,  $ID_{i+1}$ , ... mit verschiedenen Zuverlässigkeitswerten zugeordnet werden.

Erfindungsgemäß wird nunmehr nicht nur ein unbekanntes Audioobjekt am Eingang genau einem Referenz-Audioobjekt in der Referenz-Datenbank zugeordnet, und zwar nur für einen Zeitpunkt  $\Delta t$ , sondern es wird kontinuierlich gearbeitet, ohne Unterbrechung des Datenstroms am Eingang. Erfindungsgemäß wird eine Zuordnung verschiedener Ausschnitte aus Audio-Objekten jeweils zu den korrekten Audio-Objekten aus der Referenzdatenbank vorgenommen. Damit wird eine lückenlose Abfolge, d. h. ein Protokoll, der identifizierten Audio-Objekte am Eingang erhalten.

Nachfolgend wird anhand der Fig. 4a bis 5d eine besondere Schwierigkeit der kontinuierlichen Analyse eines kontinuierlichen Audiodatenstroms dargestellt. Es muß nämlich das Audio-Objekt in Abschnitte der Länge  $\Delta t_x$ , also in einzelne Blöcke zerlegt werden, um für den Teilabschnitt des Audiodatenstroms eine Zuordnung zu einem Referenzelement in der Datenbank vornehmen zu können. Es kann nun vorkommen, daß diese Zuordnung eines einzelnen Abschnitts des Audiodatenstroms nicht immer eindeutig ist und erst im Zusammenhang mit vorherigen und folgenden Zuordnungen eindeutig wird. Wenn man Einzelzuordnungen vornimmt und diese in einem weiteren Schritt lediglich zusammenfaßt, erhält man fehlerhafte Erkennungsprotokolle, was nachfolgend dargestellt wird.

Fig. 4a stellt eine lange Version eines Musikstücks XY dar, das auch durch einen in Fig. 4a dargestellten langen Fin-

gerprint dargestellt wird, wobei diesem Fingerprint das Identifikationsergebnis ID108 zugeordnet ist. Fig. 4b zeigt dasselbe für eine Kurzversion des gleichen Musikstücks XY. ID109 verweist somit auf eine Kurzversion des Musikstücks XY, während ID108 auf eine Langversion dieses Musikstücks verweist. Nachdem die Kurzversion kürzer als die Langversion ist, ist auch der Fingerprint in Fig. 4b kürzer als der Fingerprint in Fig. 4a. In der Weise, in der beide Blöcke untereinander dargestellt sind, enthalten die Musikstücke und damit auch die Fingerprints ID108 und ID109 identisches Audiomaterial bzw. identische Fingerprintdaten. ID109 ist also eine Teilmenge von ID108. Aus Fig. 4c ist damit ersichtlich, daß die lange Version einen Anfangsabschnitt hat im Zeitraum  $\Delta t_0$ , der in der kurzen Version nicht vorhanden ist. Im Mittelabschnitt zwischen  $t_1$  bis  $t_5$  sind die Langversion und die Kurzversion identisch, während die lange Version wieder zwischen dem Zeitpunkt  $t_5$  und  $t_7$  einen Musikabschnitt hat, der in der Kurzversion, die durch ID109 identifiziert ist, nicht vorhanden ist.

20

Nachfolgend wird anhand der Fig. 5a bis 5d dargestellt, wie bei einfachem Zusammenfassen, also ohne Hypothesenbildung mit den Einzelidentifikationen fehlerhafte Erkennungsprotokolle entstehen können. Es sei angenommen, daß am Eingang des Systems zum Zeitpunkt  $t_0$  das Musikstück ID108 aufgenommen wird. Ferner sei die Datenbank dahingehend wirksam, daß sie die in Fig. 5a gezeigten Elemente für die Zeitabschnitte  $\Delta t_x$  identifiziert. Es sei darauf hingewiesen, daß die Identifikation in Fig. 5a prinzipiell richtig ist, obgleich jedoch in den Zeitabschnitten  $\Delta t_1$  bis  $\Delta t_4$  sowohl ID108 als auch ID109 ausgegeben werden könnte. Letztendlich ist die Bestimmung der Identifikationsergebnisse in diesen Bereichen zweideutig, da die Datenbank bei nicht vorhandener

30

Störung sowohl ID109 als auch ID108 ausgehen wird und sich aufgrund rechnerischer Unterschiede z. B. immer für den wahrscheinlichsten Wert entscheiden wird, so daß immer aufgrund irgendwelchen Rauschens einer der beiden Identifikationsergebnisse ID108 oder ID109 ein leicht höheres Zuverlässigkeitsmaß haben wird. Im Erkennungsprotokoll, das in Fig. 5b dargestellt ist, wird somit eine falsche Identifikation dahingehend vorgenommen, daß zu keinem Zeitpunkt das durch ID109 identifizierte Stück gespielt worden ist, sondern daß lediglich das mit ID108 identifizierte Stück abgespielt worden ist.

Nachfolgend wird anhand der Fig. 5c und 5d eine weitere Alternative dargestellt. Es wird davon ausgegangen, daß die Datenbank die in Fig. 5c gezeigte Situation ausgibt. Im Erkennungsprotokoll wird wieder eine falsche Zusammenfassung gegeben, nämlich dahingehend, daß ID109 zwischen T1 und T5 vorhanden war, während dies natürlich nicht zutreffend ist. Statt dessen wurde die Langversion des Musikstücks, also ID108 von  $t_0$  bis  $t_7$  abgespielt.

Ferner sind weitere Falscherkennungsprotokolle denkbar, welche durch die Mehrdeutigkeit der Einzelerkennungen für einen Abschnitt des Audiodatenstroms im Zeitraum  $\Delta t_x$  entstehen werden.

Erfindungsgemäß wird nunmehr auf das in Fig. 6 dargestellte allgemeine Konzept zugegriffen, bei dem die für einen Zeitabschnitt  $\Delta t_x$  erhaltenen Erkennungsergebnisse, also die Ausgangssignale der Einrichtung 12 von Fig. 1, die je nach Implementierung die Einrichtungen 900, 904, 902 vereinigen kann, einem Postprocessing unterzogen werden, das im wesentlichen in der Einrichtung zum Bilden von wenigstens

zwei Hypothesen und der Einrichtung zum Untersuchen der Hypothesen von Fig. 1 entspricht. Hierauf wird dann unter Verwendung des Postprocessings, also unter Verwendung der bei dem Postprocessing erhaltenen Untersuchungsergebnisse  
5 eine Aussage über das Informationssignal in Form einer Erkennungsabfolge bzw. eines Erkennungsprotokolls gezogen.

In der Postprocessing-Stufe wird die Wahrscheinlichkeit für den Übergang von einem identifizierten Referenzaudioobjekt  
10 für den Zeitabschnitt  $\Delta t_x$  zu einem beliebigen anderen Referenzaudioobjekt für den Zeitabschnitt  $\Delta t_{x+1}$  als gleich angenommen. Hieraus werden verschiedene zunächst parallel betrachtete Hypothesen für zusammenhängende Audioabschnitte aus den Einzelerkennungen gebildet. Es ist zu beachten, daß  
15 Einzelerkennungen dann zu einer Hypothese zusammengefügt werden, wenn sie sich auf ein und dasselbe Referenzaudio-signal beziehen und zeitkontinuierlich zusammenhängen. Das Erkennungsprotokoll ergibt sich aus einem Zusammenfügen der jeweils wahrscheinlichsten Hypothesen unter Beachtung des  
20 zeitlichen Fortschritts. Nachfolgend wird ein bevorzugter Algorithmus detailliert dargestellt.

Zunächst werden verschiedene Hypothesen für zusammenhängende Audioabschnitte aus den Einzelerkennungen für die Zeit-  
25 abschnitte  $\Delta t_x$  (mit  $x = N, N+1, N+2, \dots$ ; wobei  $t_N$  der Startzeitpunkt für die jeweilige Hypothese darstellt) für je ein erkanntes Referenzaudioobjekt gebildet.

Einzelerkennungen werden dann zu einer Hypothese zusammen-  
30 gesetzt, wenn die Einzelerkennungen zeitlich kontinuierlich aufeinander folgen.

Die zeitliche Kontinuität ist ein weiteres Element, das dazu dient, festzustellen, ob eine bereits bestehende Hypothese fortgesetzt wird, oder ob eine neue Hypothese begonnen wird. So sei das Szenario betrachtet, indem ein bestimmtes z. B. Gitarrensolo in einem Stück in der Kurzversion des Stücks ziemlich am Anfang des Stücks kommt und in einer Langversion des Stücks eher in der Mitte des Stücks.

Bei einem bevorzugten Ausführungsbeispiel gibt die Datenbank, also die Einrichtung zum Bereitstellen von Identifikationsergebnissen, nicht nur eine Fingerabdruckidentifikation aus, sondern auch einen Zeitwert, der sich daraus ergibt, dass der Identifikationsfingerabdruck in der Datenbank eine Länge hat, und der eingegebene (kurze) Fingerabdruck lediglich mit einem Teil des (langen) Fingerabdrucks in der Datenbank übereinstimmt.

Bei dem vorstehend beschriebenen Szenario würde die Datenbank für das Gitarrensolo vielleicht zwei ID-Ergebnisse liefern (kurze Version und lange Version), jedoch mit zwei unterschiedlichen Zeitindizes. Der Zeitindex für das ID-Ergebnis für die kurz Version ist dabei kleiner als der Zeitindex für die lange Version. Auf der Basis des Zeitindex ist die Einrichtung zum Bilden der Hypothesen nunmehr in der Lage, Hypothesen fortzuführen (wenn eine zeitliche Kontinuität zwischen dem Zeitindex und dem letzten Zeitindex in der Hypothese ist), oder neue Hypothesen zu beginnen, wenn keine Kontinuität in dem aktuell erhaltenen Zeitindex und einem letzten Zeitindex einer Hypothese vorhanden ist.

Jede zeitliche Diskontinuität bezogen auf ein Referenzaudioobjekt erzeugt eine neue Hypothese, wenn das folgende

Element zeitlich einen größeren Abstand als ein festzulegender Zeitabstand  $T_a$  ist, oder wenn das folgende Element zeitlich vor dem vorherigen liegt.

- 5 Zur Hypothesenuntersuchung wird für jede Hypothese eine Addition der Konfidenzmaße, also der Zuverlässigkeitswerte bzw. der Maße für die Plausibilität, der Einzelerkennungen vorgenommen.
- 10 Beginnend mit dem Zeitabschnitt  $\Delta t_0$  wird dann die Hypothese mit dem höchsten Konfindenzmaß als wahr gewertet und in das Erkennungsprotokoll übernommen. Für den nächsten Zeitabschnitt, welcher der ersten Hypothese folgt, wird erneut die Hypothese mit dem höchsten Konfidenzmaß als wahr gewertet und in das Erkennungsprotokoll übernommen usw.
- 15

Für das oben bezeichnete Beispiel ergibt sich somit ein Ablauf, der anhand der Fig. 7a bis 7c dargestellt ist. Für den Zeitabschnitt  $\Delta t_0$  liefert die Datenbank, wie sie beispielsweise in Fig. 2 dargestellt ist, nur ein Identifikationsergebnis, nämlich ID108, welches eine Wahrscheinlichkeit bzw. ein Zuverlässigkeitsmaß hat, das oberhalb einer Schwelle ist. Im Zeitintervall  $\Delta t_1$ , also für den Block von Informationseinheiten, die sich über das Zeitintervall  $\Delta t_1$  erstrecken, liefert die Datenbank zwei Ergebnisse, die ein Zuverlässigkeitsmaß haben, das über einer Schwelle ist. Die zwei Ergebnisse werden auch für die Blöcke zwischen den Zeitpunkten  $t_2$  bis  $t_5$  erhalten. Dann, für den Zeitraum  $t_5$  bis  $t_7$  liefert die Datenbank wiederum lediglich ein einziges Identifikationsergebnis, dessen Zuverlässigkeitsmaß oberhalb einer Schwelle ist.

20

25

30

Die Einrichtung 14 (Fig. 1) zum Bilden von wenigstens zwei Hypothesen ist ausgebildet, um zum Zeitpunkt  $t_0$  aufgrund des Identifikationsergebnisses ID108 eine erste Hypothese zu starten, und um zum Zeitpunkt  $t_1$  aufgrund des neu hinzugekommenen Identifikationsergebnisses ID109 eine neue Hypothese, nämlich die Hypothese H2, zu starten.

Irgendwann nach dem Zeitpunkt  $t_7$  wird dann die in Fig. 7a gezeigte Hypothesensituation mit den Hypothesen H1 und H2 betrachtet, um dann aufgrund der Untersuchung der Hypothesen, die so ablaufen kann, wie es in Fig. 7b dargestellt ist, für jede Hypothese die Funktionen für die Konfidenzmaße der Einzelerkennungen, also  $x_{H1}$  und  $x_{H2}$  zu berechnen.

Wenn davon ausgegangen wird, daß zwischen  $t_1$  und  $t_5$  die Identifikationsergebnisse ID108 und ID109 mit gleicher Wahrscheinlichkeit auftreten, so wird bei dem in Fig. 7a gezeigten Ausführungsbeispiel lediglich die erste Hypothese H1 das Rennen machen, da die Hypothese zwar zwischen  $t_1$  und  $t_5$  genauso wahrscheinlich war wie die Hypothese H2, da die Hypothese H1 jedoch im Zeitraum  $\Delta t_0$  und im Zeitraum  $\Delta t_5$  und im Zeitraum  $\Delta t_6$  gilt, also ein Zuverlässigkeitsmaß für eine Einzelerkennung beiträgt, die für die Hypothese H2 nicht gegeben wird. Dies bedeutet somit für das Erkennungsprotokoll den in Fig. 7c gezeigten richtigen Fall, daß nämlich das mit ID108 bezeichnete Stück vom Zeitpunkt  $t_0$  bis zum Zeitpunkt  $t_7$  abgespielt worden ist.

Beginnend bei  $t_0$  wird somit die Hypothese H1 ausgewählt, da es bis  $t_7$  keine Hypothese mit einem höheren Konfidenzmaß gibt. Die Hypothese H2 wird verworfen, wobei prinzipiell sämtliche Hypothesen verworfen werden können, die parallel

zu einer anderen Hypothese existieren, die als die wahrscheinlichste ausgewählt worden ist.

5 Erfindungsgemäß wird somit genau die Sequenz, hier im Beispiel ein Element, nämlich ID108, protokolliert, welche auch tatsächlich am Audioeingang abgespielt wurde.

Es sei darauf hingewiesen, daß es für die Bestimmung des Endes einer Hypothese verschiedene Möglichkeiten gibt. So  
10 kann - unabhängig von der Hypothesensituation - ein Informationsentitätssende z. B. ermittelt werden aus dem Audiosignal selbst, wenn beispielsweise eine Pause mit einer bestimmten Minimallänge auftritt. Nachdem dieses Kriterium jedoch dann nicht wirkt, wenn zwischen zwei Informationsentitäten überblendet („Fading“) wird, oder wenn zwei Stücke  
15 derart schnell aufeinanderfolgen, daß keine merkbare Pause zu finden ist, wird es bevorzugt, aufgrund der in der Vergangenheit betrachteten Hypothesen ein Informationsentitätssende zu bestimmen. Dies kann beispielsweise dahingehend  
20 stattfinden, daß eine Hypothese dann als beendet angesehen wird, wenn z. B. zwei oder mehr Blöcke kein Identifikationsergebnis mehr mit einem Zuverlässigkeitswert oberhalb einer bestimmten Minimalschwelle der Einrichtung 14 zum Bilden von Hypothesen zugeführt werden. Alternativ kann  
25 auch, beispielsweise für den in Fig. 3 gezeigten Fall einfach damit begonnen werden, zu irgendeinem Zeitpunkt in die Vergangenheit gerichtet die Werte der Hypothesen für eine vorbestimmte Anzahl von Blöcken aufzuzaddieren, um dann zu sehen, welche Hypothese am Ende, also nach einer bestimmten  
30 Anzahl von z. B. 20 Blöcken für bestimmte Blöcke den höchsten Wert hatte und damit überlebt hat und die anderen Hypothesen „ausgestochen“ hat. Dies würde bei dem in Fig. 3 gezeigten Beispiel bedeuten, daß die Hypothesen, daß die In-

formationsentität ID1 oder ID2 oder ID3 ist, auch für die Zeitabschnitte  $\Delta t_7$  und  $\Delta t_8$  weitergeführt werden würden, wobei dies jedoch nichts an der Erkennung von ID1 ändern würde, da neue Hypothesen, also die Hypothese für ID108, ID109, ID4 und ID8 erst wesentlich später, also für Blöcke von  $\Delta t_7$  und  $\Delta t_8$  oder darüber begonnen werden und damit erst viel später oder überhaupt nicht derart hohe kombinierte Zuverlässigkeitswerte erreichen.

10 Aus dem vorstehenden wird ersichtlich, daß das Ende einer Hypothese nicht unbedingt aktiv bestimmt werden muß, sondern daß sich dieses Ende aus der Analyse der Vergangenheit, also der begonnenen Hypothesen automatisch ergeben kann. Vorzugsweise wird daher immer dann, wenn ein neues  
15 Identifikationsergebnis mit einem Zuverlässigkeitsmaß oberhalb einer Signifikanzschwelle in Erscheinung tritt, eine neue Hypothese begonnen, wobei dann irgendwann in die Vergangenheit geblickt wird, um zu sehen, welche Hypothese für einen bestimmten Zeitraum überlebt, wobei hierfür nicht ex-  
20 plizit ein Ende einer Hypothese bestimmt werden muß, da es sich automatisch selbst ergibt.

Abhängig von der Gegebenheit, kann das erfindungsgemäße Verfahren in Hardware oder in Software implementiert werden. Die Implementierung kann auf einem digitalen Speichermedium, insbesondere einer Diskette oder CD mit elektronisch auslesbaren Steuersignalen erfolgen, die so mit einem programmierbaren Computersystem zusammenwirken können, dass das Verfahren ausgeführt wird. Allgemein besteht die Erfindung somit auch in einem Computer-Programm-Produkt mit einem auf einem maschinenlesbaren Träger gespeicherten Programmcode zur Durchführung des erfindungsgemäßen Verfahrens, wenn das Computer-Programm-Produkt auf einem Rechner

abläuft. In anderen Worten ausgedrückt, kann die Erfindung somit als ein Computer-Programm mit einem Programmcode zur Durchführung des Verfahrens realisiert werden, wenn das Computer-Programm auf einem Computer abläuft.

Patentansprüche

1. Vorrichtung zum Analysieren eines Informationssignals,  
das eine Folge (802) von Blöcken von Informationsein-  
heiten (804) aufweist, wobei eine Mehrzahl von aufein-  
5 anderfolgenden Blöcken der Folge von Blöcken eine In-  
formationsentität (806) darstellt, unter Verwendung  
einer Folge von Fingerabdrücken (FAi) für die Folge  
von Blöcken, so daß die Folge von Blöcken durch die  
10 Folge von Fingerabdrücken dargestellt ist, mit folgen-  
den Merkmalen:

einer Einrichtung (12) zum Bereitstellen von Identifi-  
kationsergebnissen (IDi) für aufeinanderfolgende Fin-  
15 gerabdrücke, wobei ein Identifikationsergebnis eine  
Zugehörigkeit eines Blocks von Informationseinheiten  
zu einer vorbestimmten Informationsentität darstellt,  
und wobei es für jedes Identifikationsergebnis ein Zu-  
verlässigkeitsmaß gibt, wobei die Einrichtung (12) zum  
20 Bereitstellen ausgebildet ist, um für einen ersten  
Fingerabdruck ein erstes Identifikationsergebnis zu  
erzeugen, und um für einen folgenden Block ein zweites  
Identifikationsergebnis zu erzeugen, das sich von dem  
ersten Identifikationsergebnis unterscheidet;

25 einer Einrichtung (14) zum Bilden von wenigstens zwei  
Hypothesen aus den Identifikationsergebnissen für die  
aufeinanderfolgenden Fingerabdrücke, wobei eine erste  
Hypothese eine Annahme für die Zugehörigkeit der Folge  
30 von Blöcken zu einer ersten Informationsentität ist,  
und wobei eine zweite Hypothese eine Annahme für die  
Zugehörigkeit der Folge von Blöcken zu einer zweiten  
Informationsentität ist, wobei die Einrichtung (14)  
zum Bilden ausgebildet ist, um ansprechend auf das  
35 erste Identifikationsergebnis die erste Hypothese zu  
beginnen oder die bereits bestehende erste Hypothese  
weiterzuführen und um ansprechend auf das zweite Iden-  
tifikationsergebnis die zweite Hypothese zu beginnen

oder die bereits bestehende zweite Hypothese weiterzuführen;

5 einer Einrichtung (16) zum Untersuchen der wenigstens zwei Hypothesen durch Zusammenfassen der Zuverlässigkeitsmaße der Hypothesen, um ein Untersuchungsergebnis (18) zu erhalten; und

10 einer Einrichtung (20) zum Treffen einer Aussage über das Informationssignal basierend auf dem Untersuchungsergebnis.

2. Vorrichtung nach Anspruch 1, bei der die Einrichtung (16) zum Untersuchen ausgebildet ist, um die Hypothesen im Hinblick auf für die Hypothesen gültige Wahrscheinlichkeitsinformationen zu untersuchen.

3. Vorrichtung nach Anspruch 1 oder 2, bei der die Einrichtung (20) zum Treffen einer Aussage ausgebildet ist, um zu bestimmen, daß die Folge von Blöcken eine Informationsentität darstellt, die eine Hypothese hat, die am wahrscheinlichsten ist, oder daß eine Informationsentität mit dem Fingerabdruck beendet ist, der als zeitlich letzter zu der wahrscheinlichsten Hypothese beiträgt, oder daß eine Informationsentität in dem Informationssignal vorhanden ist oder nicht.

4. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Einrichtung (12) zum Bereitstellen ausgebildet ist, um für einen Fingerabdruck zwei unterschiedliche Identifikationsergebnisse zu erzeugen.

5. Vorrichtung nach Anspruch 4, bei der die Einrichtung (12) zum Bereitstellen ausgebildet ist, um für jeden der zwei unterschiedlichen Identifikationsergebnisse ein Zuverlässigkeitsmaß zu erzeugen.

6. Vorrichtung nach Anspruch 4 oder 5, bei der die Einrichtung (14) zum Bilden ausgebildet ist, um der ersten Hypothese ein erstes der zwei Identifikationsergebnisse und der zweiten Hypothese ein zweites der zwei Identifikationsergebnisse zuzuordnen.  
5
7. Vorrichtung nach einem der Ansprüche 3 bis 6, bei der die Einrichtung (16) zum Untersuchen ausgebildet ist, um die Hypothese zu bestimmen, die ein höheres zusammengefaßtes Zuverlässigkeitsmaß hat.  
10
8. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Einrichtung (14) zum Bilden ausgebildet ist, um die erste oder zweite Hypothese zu beenden, wenn eine vorbestimmte Anzahl von Blöcken weder ein Identifikationsergebnis, das auf die erste Informationse-  
15  
ntität hinweist, noch ein Identifikationsergebnis, das auf die zweite Informationse-  
ntität hinweist, erhalten wird.  
20
9. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Einrichtung (14) zum Bilden ausgebildet ist, um die erste oder zweite Hypothese zu beenden, wenn ein detektiertes Ereignis in dem Informationssig-  
25  
nal auftritt.
10. Vorrichtung nach Anspruch 9, bei der ein Ereignisdetektor vorhanden ist, der ausgebildet ist, um einen Energiepegel in einem Block von Informationseinheiten, der unter einem Schwellenpegel liegt, als das Ereignis zu detektieren.  
30
11. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Einrichtung (12) zum Bereitstellen ausgebildet ist, um für jeden Fingerabdruck nur das zuverlässigste Identifikationsergebnis ohne oder mit Zuverlässigkeitsmaß auszugeben, um für einen Fingerabdruck eine vorbestimmte Anzahl von zuverlässigsten Fingerab-  
35

drücken jeweils mit oder ohne Zuverlässigkeitsmaß auszugeben, oder um für einen Fingerabdruck lediglich die Identifikationsergebnisse mit oder ohne Zuverlässigkeitsmaße auszugeben, die ein Zuverlässigkeitsmaß oberhalb einer Schwelle haben.

5  
12. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Einrichtung (16) zum Untersuchen ausgebildet ist, um explizite oder implizite Zuverlässigkeitsmaße, die zu einer Hypothese gehören, zusammenzuaddieren, um ein zusammengefaßtes Zuverlässigkeitsmaß zu erhalten.

10  
13. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Einrichtung (12) zum Bereitstellen ausgebildet ist,

um mit einem Fingerabdruck eine Suche in einer Datenbank, in der Fingerabdrücke von Referenz-Informationsentitäten gespeichert sind, durchzuführen, und

20  
um eine Anzahl von Identifikationsergebnissen sowie ein Abstandsmaß für jedes Identifikationsergebnis als Hinweis auf ein Zuverlässigkeitsmaß für jedes Identifikationsergebnis zu liefern.

25  
14. Vorrichtung nach Anspruch 13, bei der die Einrichtung (12) zum Bereitstellen ausgebildet ist, um für jedes Identifikationsergebnis, für das nicht bereits eine Hypothese existiert, eine neue Hypothese zu beginnen, wenn ein Abstandsmaß für das Identifikationsergebnis eine Beziehung zu einer Schwelle hat, die auf einen kleineren Abstand als einen Schwellenabstand hinweist.

30  
35  
15. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Einrichtung (16) zum Untersuchen ausgebildet ist, um ansprechend auf eine Bestimmung alle Hypo-

thesen für die aufeinanderfolgenden Fingerabdrücke zu beenden, die für die Fingerabdrücke gebildet worden sind, die von der wahrscheinlichsten Hypothese erfaßt sind.

5

16. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der das Informationssignal ein Audiosignal umfaßt, bei der die Informationseinheit Audioabtastwerte im Zeit- oder Frequenzbereich sind, und bei der eine Informationsentität ein Musikstück, eine gesprochene Sequenz oder einen Geräuschabschnitt umfaßt.

10

17. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der ein Fingerabdruck für einen Block durch eine Zeit-Frequenz-Umsetzung und/oder durch Berechnung eines spektralen Flachheitsmaßes für ein Ergebnis der Zeit-Frequenz-Umsetzung ermittelt wird.

15

18. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der ein Fingerabdruck für einen Block so erzeugt wird, daß der Fingerabdruck eine Datenmenge hat, die kleiner ist als eine Datenmenge des Blocks.

20

19. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Einrichtung (12) zum Bereitstellen von Identifikationsergebnissen ausgebildet ist, um neben einem Identifikationsergebnis auch einen neuen Zeitindex für das Identifikationsergebnis zu liefern, und

25

30 bei der die Einrichtung (14) zum Bilden von Hypothesen ausgebildet ist, um eine Hypothese weiterzuführen, wenn eine Kontinuität zwischen einem aktuellsten Zeitindex in der Hypothese und dem neuen Zeitindex existiert, oder um eine Hypothese zu starten, wenn die Kontinuität nicht existiert.

30

35

20. Verfahren zum Analysieren eines Informationssignals, das eine Folge (802) von Blöcken von Informationseinheiten (804) aufweist, wobei eine Mehrzahl von aufeinanderfolgenden Blöcken der Folge von Blöcken eine Informationsentität (806) darstellt, unter Verwendung einer Folge von Fingerabdrücken (FAi) für die Folge von Blöcken, so daß die Folge von Blöcken durch die Folge von Fingerabdrücken dargestellt ist, mit folgenden Schritten:

10

Bereitstellen (12) von Identifikationsergebnissen (IDi) für aufeinanderfolgende Fingerabdrücke, wobei ein Identifikationsergebnis eine Zugehörigkeit eines Blocks von Informationseinheiten zu einer vorbestimmten Informationsentität darstellt, und wobei es für jedes Identifikationsergebnis ein Zuverlässigkeitsmaß gibt, wobei im Schritt des Bereitstellens für einen ersten Fingerabdruck ein erstes Identifikationsergebnis und für einen folgenden Block ein zweites Identifikationsergebnis erzeugt werden, das sich von dem ersten Identifikationsergebnis unterscheidet;

15

20

25

30

Bilden (14) von wenigstens zwei Hypothesen aus den Identifikationsergebnissen für die aufeinanderfolgenden Fingerabdrücke, wobei eine erste Hypothese eine Annahme für die Zugehörigkeit der Folge von Blöcken zu einer ersten Informationsentität ist, und wobei die zweite Hypothese eine Annahme für eine Zugehörigkeit der Folge von Blöcken zu einer zweiten Informationsentität ist, wobei der Schritt des Bildens folgende Schritte aufweist:

35

ansprechend auf das erste Identifikationsergebnis, Beginnen der ersten Hypothese oder Weiterführen der bereits bestehenden ersten Hypothese, und, ansprechend auf das zweite Identifikationsergebnis, Beginnen der zweiten Hypothese oder

Weiterführen der bereits bestehenden zweiten Hypothese;

5           Untersuchen (16) der wenigstens zwei Hypothesen durch  
Zusammenfassen der Zuverlässigkeitsmaße der Hypothe-  
sen, um ein Untersuchungsergebnis (18) zu erhalten;  
und

10           Treffen (20) einer Aussage über das Informationssignal  
basierend auf dem Untersuchungsergebnis.

21.   Computerprogramm mit einem Programmcode zum Durchfüh-  
ren eines Verfahrens gemäß Patentanspruch 20, wenn das  
15           Programm auf einem Computer abläuft.

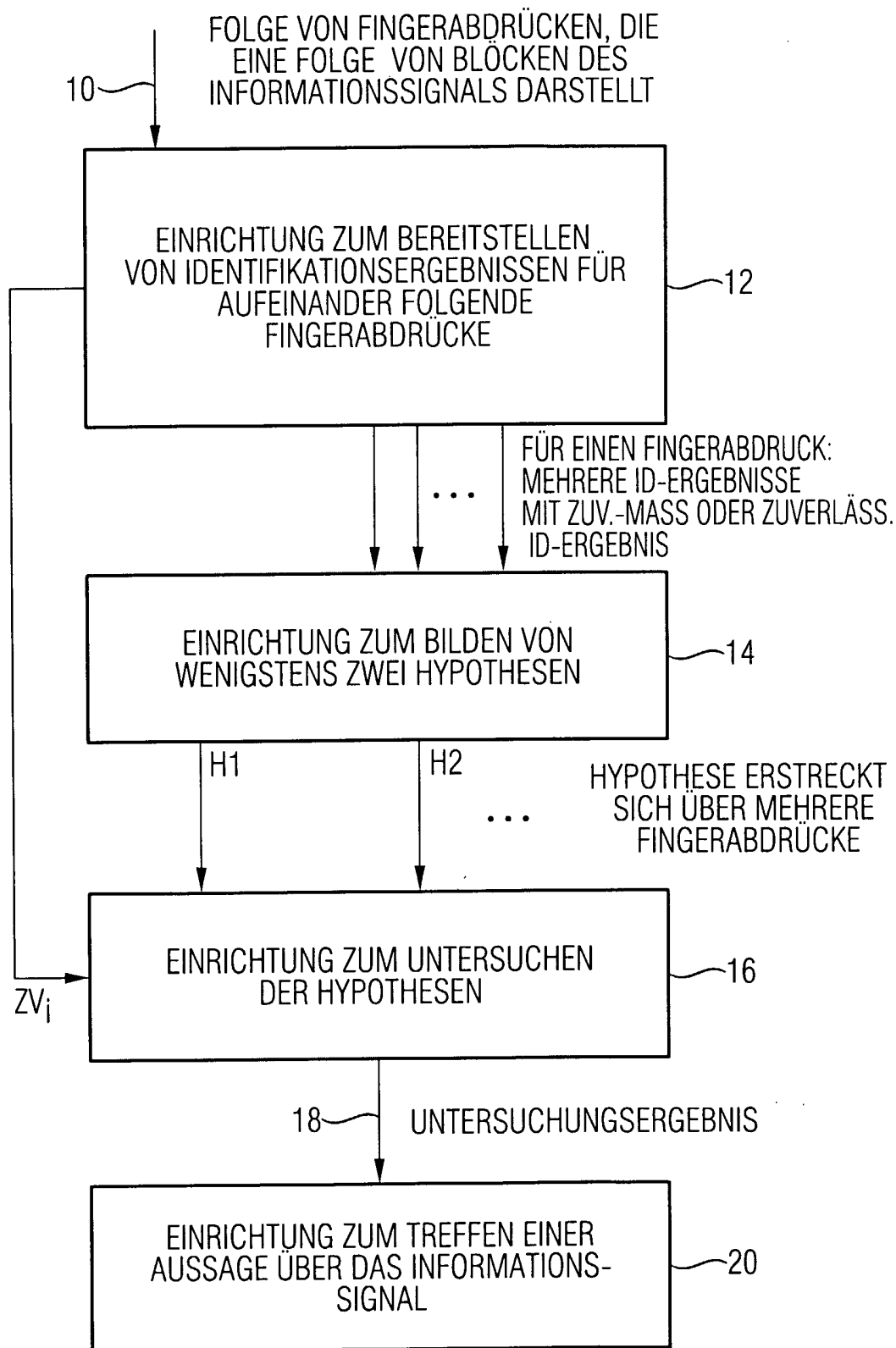
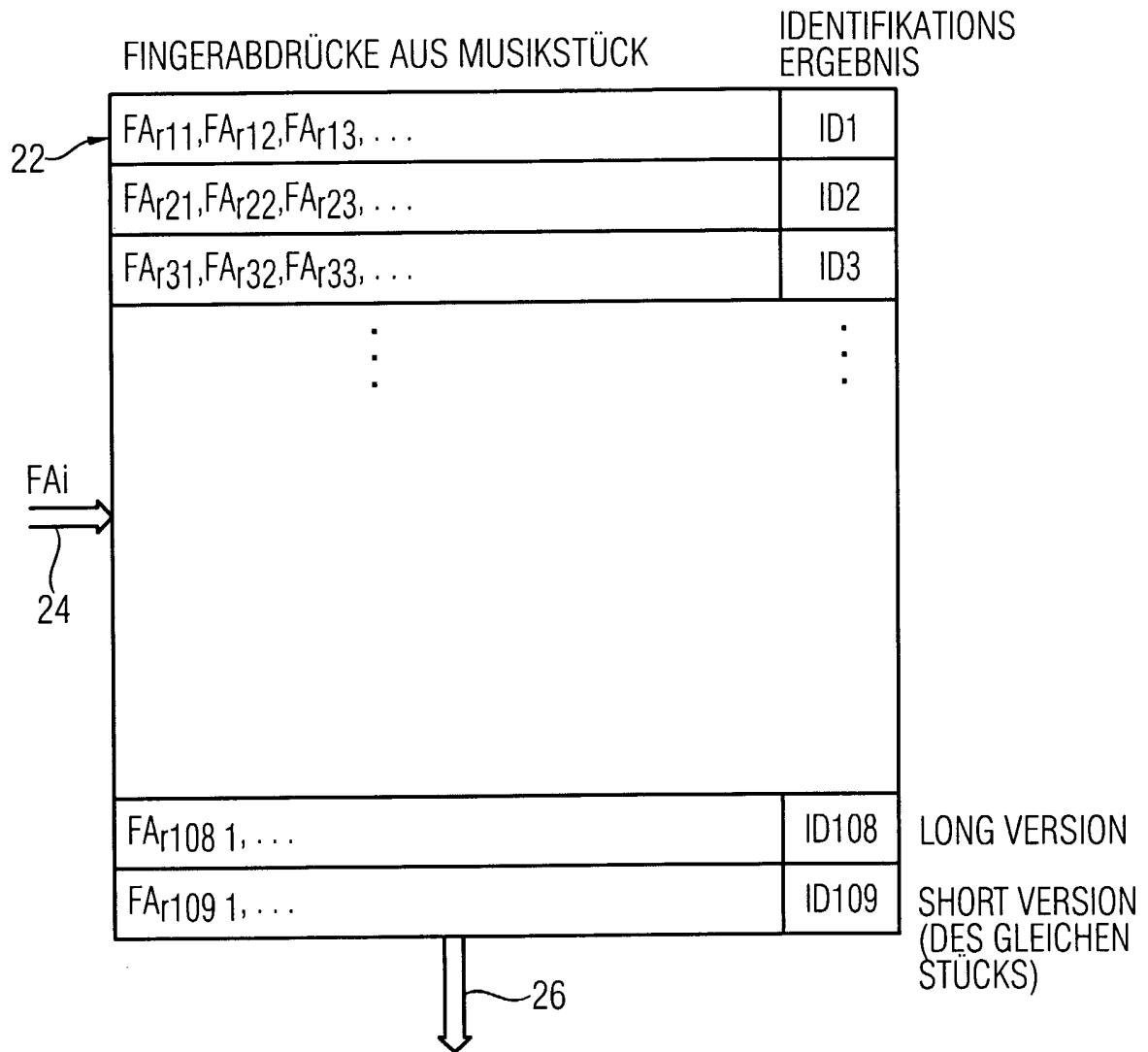


FIG. 1

DATENBANK



ID <sub>x</sub>	ZV <sub>1</sub>	(60%)
ID <sub>y</sub>	ZV <sub>2</sub>	(50%)
ID <sub>z</sub>	ZV <sub>3</sub>	(40%)

28 →

ZV<sub>n</sub>: ZUVERLÄSSIGHEITSMASS

FIG. 2

3 / 7

	$\Delta t_1$	$\Delta t_2$	$\Delta t_3$	$\Delta t_4$	$\Delta t_5$	$\Delta t_6$	$\Delta t_7$	$\Delta t_8$
	FA1	FA2	FA3	FA4	FA5	FA6	FA7	FA8
	ID1 40%	ID1 60	ID1 80	ID1 60	ID1 90	ID1 90	ID108 90	ID108 80
	ID2 60	ID2 70	ID2 30	ID2 40	ID2 20	ID2 10	ID109 85	ID109 70
	ID3 30	ID3 20	ID3 30	ID3 40	ID3 10	ID3 5	ID4 10	ID4 10
H1	ID1	ID1	ID1	ID1	ID1	ID1	ID108	ID108
H2	ID2	ID2	ID2	ID2	ID2	ID2	ID109	ID109
H3	ID3	ID3	ID3	ID3	ID3	ID3	ID4	ID4
H4								ID8

FÜR  $\Delta t_1$ - $\Delta t_6$

$\Sigma$  ID1: 420

$\Sigma$  ID2: 230

$\Sigma$  ID3: 135

MUSIKSTÜCK-  
ÜBERGANG

⇒ AUSSAGE: -INFORMATIONSENTITÄT VON  $\Delta t_1$ - $\Delta t_6$  IST MUSIKSTÜCK ID1; H1 WAR RICHTIG; UND/ODER  
 - $\Delta t_6$  IST ENDE DER 1. INFORMATIONSENTITÄT BZW.  $\Delta t_7$  IST BEGINN DER 2. INFORMATIONSENTITÄT; UND/ODER  
 -ID1 IST IN INFORMATIONSSIGNAL ENTHALTEN

FIG. 3

4 / 7

ID 108 IST EINE LANGE VERSION EINES MUSIKSTÜCKES XY (LÄNGE Z.B. 7:32, SYMBOLISCH DURCH FOLGENDEN FINGERPRINT DARGESTELLT:



FIG. 4A

ID109 IST EINE KURZVERSION DES GLEICHEN MUSIKSTÜCKES XY (Z.B. EINE SOG. "RADIO-VERSION" ODER "RADIO-EDIT", LÄNGE Z.B. 4:38), SYMBOLISCH DURCH FOLGENDEN FINGERPRINT DARGESTELLT:

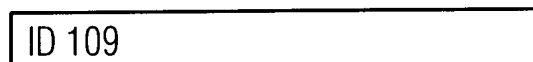


FIG. 4B

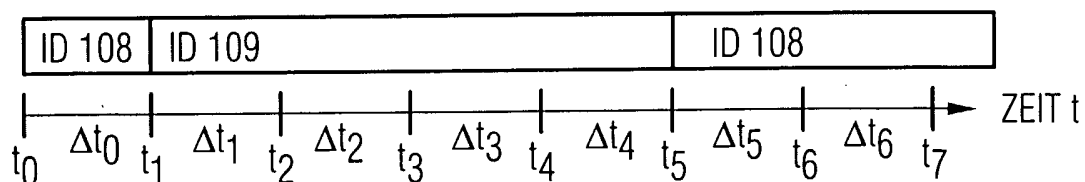


FIG. 4C

5 / 7

IDENTIFIZIERTE ELEMENTE FÜR DIE ZEITABSCHNITTE  $\Delta t_x$  (MIT  $x=0,1,2,3,4,5,6$ )

ID 108	ID 109	ID 108	ID 109	ID 109	ID 108	ID 108
$\Delta t_0$	$\Delta t_1$	$\Delta t_2$	$\Delta t_3$	$\Delta t_4$	$\Delta t_5$	$\Delta t_6$

FIG. 5A

ZUSAMMENFASSUNG IM ERKENNUNGSPROTOKOLL:

1.  $t_0$  BIS  $t_1$ : ID 108 ( $\Delta t_0$ )
2.  $t_1$  BIS  $t_2$ : ID 109 ( $\Delta t_1$ )
3.  $t_2$  BIS  $t_3$ : ID 108 ( $\Delta t_2$ )
4.  $t_3$  BIS  $t_5$ : ID 109 ( $\Delta t_3, \Delta t_4$ )
5.  $t_5$  BIS  $t_7$ : ID 108 ( $\Delta t_5, \Delta t_6$ )

FIG. 5B

IDENTIFIZIERTE ELEMENTE FÜR DIE ZEITABSCHNITTE  $\Delta t_x$  (MIT  $x=0,1,2,3,4,5,6$ )

ID 108	ID 109	ID 109	ID 109	ID 109	ID 108	ID 108
--------	--------	--------	--------	--------	--------	--------

FIG. 5C

ZUSAMMENFASSUNG IM ERKENNUNGSPROTOKOLL:

1.  $t_0$  BIS  $t_1$ : ID 108 ( $\Delta t_0$ )
2.  $t_1$  BIS  $t_5$ : ID 109 ( $\Delta t_1 - \Delta t_4$ )
3.  $t_5$  BIS  $t_7$ : ID 108 ( $\Delta t_5 - \Delta t_6$ )

FIG. 5D

6 / 7

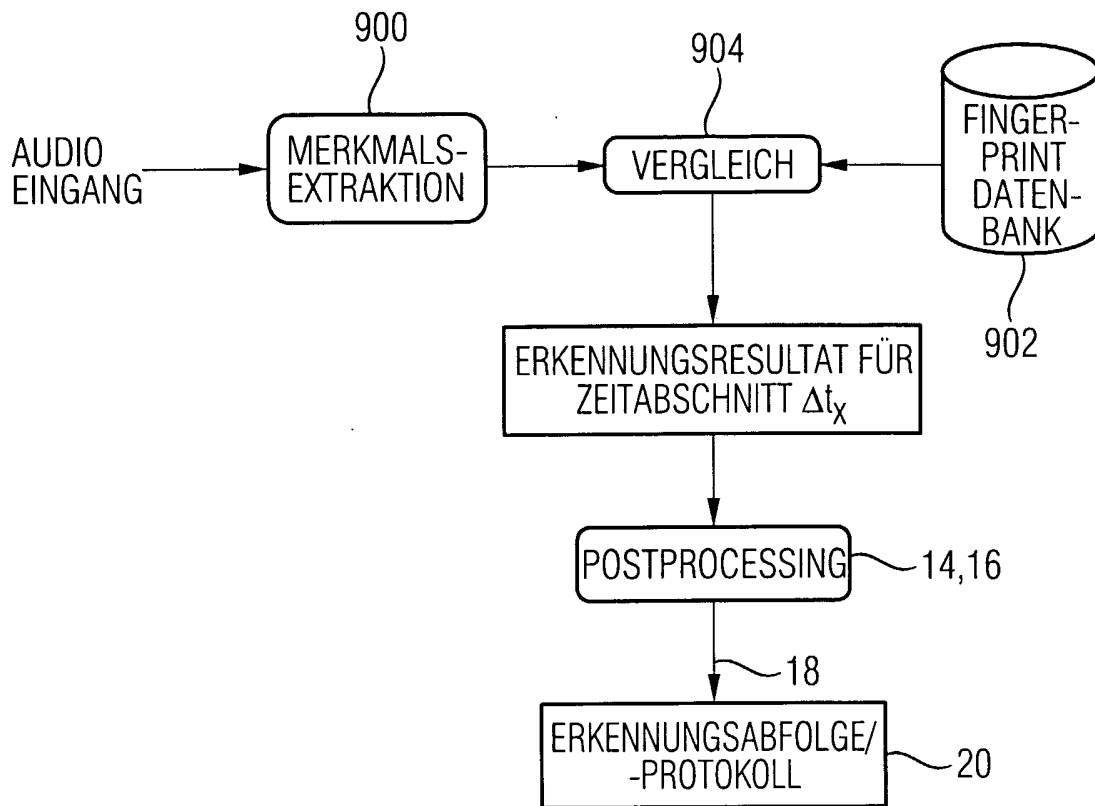


FIG 6

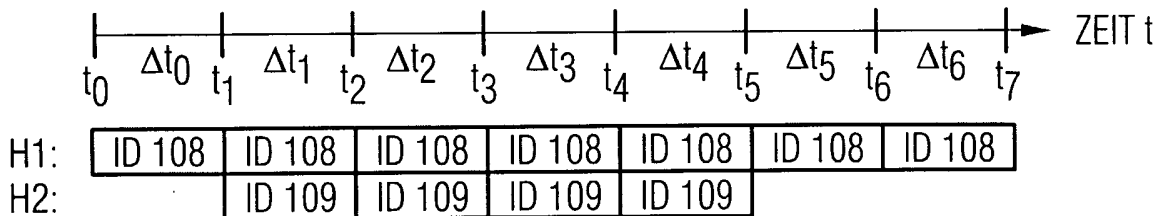


FIG 7A

HYPOTHESE 1 (H1):  $t_0$  BIS  $t_7$ : ID 108, KONFIDENZMASS  $x_{H1}\%$   
 HYPOTHESE 2 (H2):  $t_1$  BIS  $t_5$ : ID 109, KONFIDENZMASS  $x_{H2}\%$   
 MIT  $x_{H1}\% > x_{H2}\%$ , WOBEI  $x_{H1}\%$  UND  $x_{H2}\%$  JE EINE FUNKTION DER KONFIDENZMASSE DER EINZELERKENNUNGEN SIND.

FIG 7B

DIES BEDEUTET FÜR DAS ERKENNUNGSPROTOKOLL:

1.  $t_0$  BIS  $t_7$ : ID 108 ( $\Delta t_0$ - $\Delta t_6$ )

FIG 7C

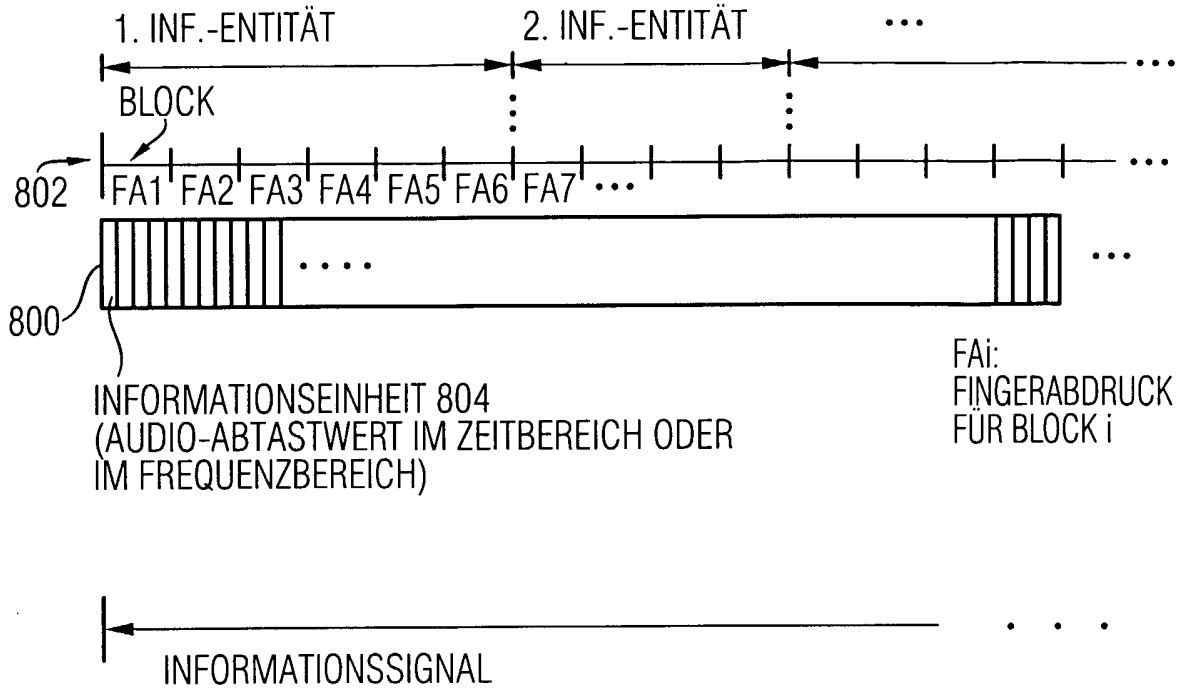


FIG. 8

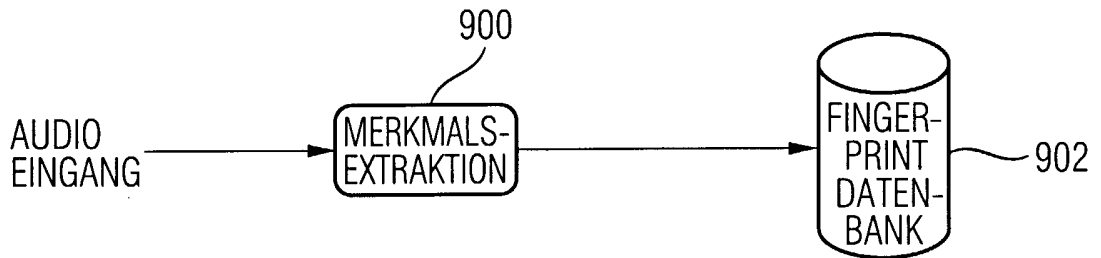


FIG. 9

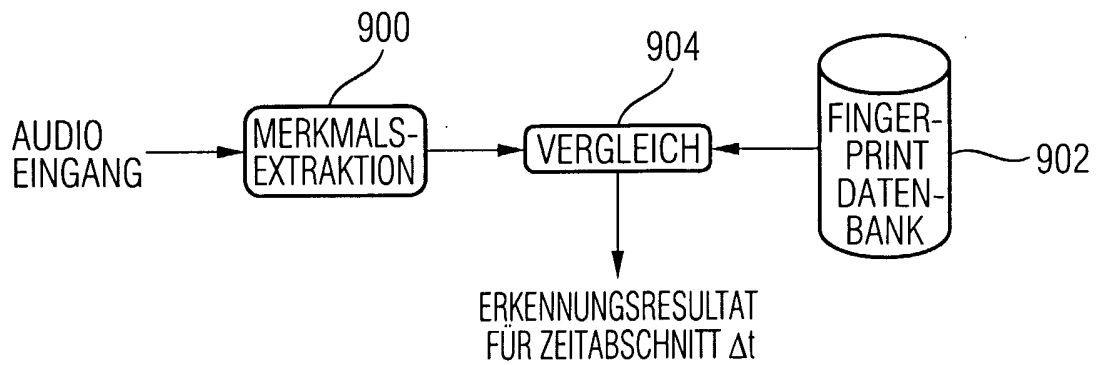


FIG. 10

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/EP2005/005004

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC 7 G10L11/00		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols) IPC 7 G10L		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, INSPEC, WPI Data, IBM-TDB, PAJ		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BATLLE E ET AL: "Automatic song identification in noisy broadcast audio" PROCEEDINGS OF THE FOURTH IASTED INTERNATIONAL CONFERENCE SIGNAL AND IMAGE PROCESSING ACTA PRESS ANAHEIM, CA, USA, 2002, pages 230-235, XPO02337265 ISBN: 0-88986-340-7 abstract page 230, right-hand column, last paragraph - page 231, left-hand column, paragraph 4 page 232, left-hand column, paragraph 3 - right-hand column, paragraph 1 ----- -/--	1-21
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C.		
<input checked="" type="checkbox"/> Patent family members are listed in annex.		
° Special categories of cited documents :		
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	
"E" earlier document but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.	
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family	
"P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search  <p style="text-align: center;">26 July 2005</p>	Date of mailing of the international search report  <p style="text-align: center;">08/08/2005</p>	
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer  <p style="text-align: center;">Greiser, N</p>	

INTERNATIONAL SEARCH REPORT

International Application No  
PCT/EP2005/005004

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>CHAI ET AL.: "Folk Music Classification Using Hidden Markov Models" PROCEEDINGS OF INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE, 'Online! June 2001 (2001-06), pages 1-6, XP002337266 Las Vegas (USA) Retrieved from the Internet: URL:http://web.media.mit.edu/{chaiwei/papers/chai_ICAI183.pdf}&gt; 'retrieved on 2005-07-21! abstract</p> <p style="text-align: center;">-----</p>	1,20
A	<p>WO 02/11123 A (SHAZAM ENTERTAINMENT LIMITED; WANG, AVERY, LI-CHUN; SMITH, JULIUS, O.) 7 February 2002 (2002-02-07) abstract</p> <p style="text-align: center;">-----</p>	1,20

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP2005/005004

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0211123 A	07-02-2002	US 2002083060 A1	27-06-2002
		AU 8976601 A	13-02-2002
		BR 0112901 A	10-06-2003
		CN 1592906 A	09-03-2005
		WO 0211123 A2	07-02-2002
		EP 1307833 A2	07-05-2003
		JP 2004505328 T	19-02-2004
		US 2004199387 A1	07-10-2004

---

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen  
PCT/EP2005/005004

<b>A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES</b> IPK 7 G10L11/00		
Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK		
<b>B. RECHERCHIERTE GEBIETE</b>		
Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole) IPK 7 G10L		
Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen		
Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe) EPO-Internal, INSPEC, WPI Data, IBM-TDB, PAJ		
<b>C. ALS WESENTLICH ANGESEHENE UNTERLAGEN</b>		
Kategorie°	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	BATLLE E ET AL: "Automatic song identification in noisy broadcast audio" PROCEEDINGS OF THE FOURTH IASTED INTERNATIONAL CONFERENCE SIGNAL AND IMAGE PROCESSING ACTA PRESS ANAHEIM, CA, USA, 2002, Seiten 230-235, XP002337265 ISBN: 0-88986-340-7 Zusammenfassung Seite 230, rechte Spalte, letzter Absatz - Seite 231, linke Spalte, Absatz 4 Seite 232, linke Spalte, Absatz 3 - rechte Spalte, Absatz 1 ----- -/--	1-21
<input checked="" type="checkbox"/>	Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen	<input checked="" type="checkbox"/> Siehe Anhang Patentfamilie
° Besondere Kategorien von angegebenen Veröffentlichungen : "A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist "E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist "L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt) "O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht "P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist		"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist "X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden "Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist "&" Veröffentlichung, die Mitglied derselben Patentfamilie ist
Datum des Abschlusses der internationalen Recherche 26. Juli 2005		Absenddatum des internationalen Recherchenberichts 08/08/2005
Name und Postanschrift der Internationalen Recherchenbehörde Europäisches Patentamt, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Bevollmächtigter Bediensteter Greiser, N

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen  
PCT/EP2005/005004

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN		
Kategorie°	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	<p>CHAI ET AL.: "Folk Music Classification Using Hidden Markov Models"                      PROCEEDINGS OF INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE, 'Online!                      Juni 2001 (2001-06), Seiten 1-6,                      XP002337266                      Las Vegas (USA)                      Gefunden im Internet:                      URL:<a href="http://web.media.mit.edu/~chaiwei/papers/chai_ICAI183.pdf">http://web.media.mit.edu/~chaiwei/papers/chai_ICAI183.pdf</a>                      'gefunden am 2005-07-21!                      Zusammenfassung                      -----</p>	1,20
A	<p>WO 02/11123 A (SHAZAM ENTERTAINMENT LIMITED; WANG, AVERY, LI-CHUN; SMITH, JULIUS, O.) 7. Februar 2002 (2002-02-07)                      Zusammenfassung                      -----</p>	1,20

**INTERNATIONALE RECHERCHENBERICHT**

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen

PCT/EP2005/005004

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
WO 0211123 A	07-02-2002	US 2002083060 A1	27-06-2002
		AU 8976601 A	13-02-2002
		BR 0112901 A	10-06-2003
		CN 1592906 A	09-03-2005
		WO 0211123 A2	07-02-2002
		EP 1307833 A2	07-05-2003
		JP 2004505328 T	19-02-2004
		US 2004199387 A1	07-10-2004
-----			